

APPLICANT: KHAN, Farrah A.
APPLICANT: YAO, Monique G.
APPLICANT: EMERLING, Brooke M.
TITLE OF INVENTION: INTRACELLULAR SIGNALING MOLECULES
FILE REFERENCE: PF-0868 USN
CURRENT APPLICATION NUMBER: US/10/451,207
CURRENT FILING DATE: 2003-06-18
PRIOR APPLICATION NUMBER: PCT/US01/50315
PRIOR FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: US 60/257,804
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: US 60/260,102
PRIOR FILING DATE: 2001-01-05
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PERL Program
SEQ ID NO 2
LENGTH: 670
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040038267A1 5560161CDI
US-10-451-207-2

Query Match 91.4%; Score 3449.5; DB 12; Length 670;
Best Local Similarity 92.7%; Pred. No. 1.9e-293;
Matches 670; Conservative 0; Mismatches 0; Indels 53; Gaps 1;
1 MILEERPDGAGAGESPRLOISRRKPRKTRVSSLRGRREGRLDVCAMGRCVHRGEDPVR 60
1 MILEERPDGAGAGESPRLO----- 20
61 VHVGPMPOLHAGVCDLSLTQCGLOSRRAQIHQOIKDELQMTGAEINLYRATSNRVR 120
21 -----GCDLSLTQCGLOSRRAQIHQOIKDELQMTGAEINLYRATSNRVR 67
121 ETVALELSYVNSNLQKLEELSGVDPGRHSGEAVTVPMIPLGLKETKELDWSTPLK 180
68 ETVALELSYVNSNLQKLEELSGVDPGRHSGEAVTVPMIPLGLKETKELDWSTPLK 127
181 ELISVHFGEDGASVFAEIRELEALRQAMRTPSRNESGLELLTAYYNQICFLDARFLTPAR 240
128 ELISVHFGEDGASVFAEIRELEALRQAMRTPSRNESGLELLTAYYNQICFLDARFLTPAR 187
241 SLGLFFHWYDLSLTGVPAAQORALAPEKGSVLFNICALHTQIGARDSCCTGRRAMEAFQ 300
188 SLGLFFHWYDLSLTGVPAAQORALAPEKGSVLFNICALHTQIGARDSCCTGRRAMEAFQ 247
301 RAAGAFSLLENFSHAPSPPMSASLQALQMAQAQECVFEGLSPFASMAPQDCCLAQL 360
248 RAAGAFSLLENFSHAPSPPMSASLQALQMAQAQECVFEGLSPFASMAPQDCCLAQL 307
361 RLAGEAQAQAAEYRLVHRTWAOVPHDYVPSWTALVHVKAEPYFRSLAHYVAMALCDGS 420
308 RLAGEAQAQAAEYRLVHRTWAOVPHDYVPSWTALVHVKAEPYFRSLAHYVAMALCDGS 367
421 PATEGELPHEQVFLQPTSSKRGVPLPOELERRQLGKAHLKRAILGQEEALRHLC 480
368 PATEGELPHEQVFLQPTSSKRGVPLPOELERRQLGKAHLKRAILGQEEALRHLC 427
481 RVLRVDLLRAVISQTLQRLSLAKYAEILDREDDFCEAAEAPDIQPKTHQKPEARMPLSQG 540
428 RVLRVDLLRAVISQTLQRLSLAKYAEILDREDDFCEAAEAPDIQPKTHQKPEARMPLSQG 487
541 KGPDIHRLGFLSVFSKRWRLVGPVHLTRGGGFLTLRGDSPVLIAAVIPGSAQAAA 600
488 KGPDIHRLGFLSVFSKRWRLVGPVHLTRGGGFLTLRGDSPVLIAAVIPGSAQAAA 547
601 GLKEGDYIVSVNGQPCRWHEAVTTELKAGAGASLOVSVLLPSRLPSLGDPRVLL 660
548 GLKEGDYIVSVNGQPCRWHEAVTTELKAGAGASLOVSVLLPSRLPSLGDPRVLL 607
661 GPRGLRSQREHGCKTFASTWASPRPLNWSRKAQQGKTGCGPQPCAPVKPAPSSILKHP 720

Db 608 GPRGLRSQREHGCKTFASTWASPRPLNWSRKAQQGKTGCGPQPCAPVKPAPSSILKHP 667
QY 721 GWP 723
Db 668 GWP 670
RESULT 4
US-10-112-944-318
Sequence 318, Application US/10112944
Publication No. US20040048249A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Yang, Yonghong
APPLICANT: Weng, Gezhi
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Xue, Aidong J.
APPLICANT: Wang, Jian-Rui
APPLICANT: Wehrman, Tom
APPLICANT: Ghosh, Malabika
APPLICANT: Wang, Dunrui
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Zhiwei
TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
TITLE OF INVENTION: Secreted Polypeptides
FILE REFERENCE: 805A
CURRENT APPLICATION NUMBER: US/10/112,944
CURRENT FILING DATE: 2002-03-28
PRIOR APPLICATION NUMBER: US 09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US 09/491,404
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: US 09/496,914
PRIOR FILING DATE: 2000-02-03
PRIOR APPLICATION NUMBER: US 09/515,126
PRIOR FILING DATE: 2000-02-28
PRIOR APPLICATION NUMBER: US 09/519,705
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: US 09/552,929
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: US 09/577,408
NUMBER OF SEQ ID NOS: 924
SOFTWARE: pt_FL_genes Version 5.0
SEQ ID NO 318
LENGTH: 669
TYPE: PRT
ORGANISM: Homo sapiens
US-10-112-944-318
Query Match 91.0%; Score 3435; DB 12; Length 669;
Best Local Similarity 92.5%; Pred. No. 3.6e-292;
Matches 669; Conservative 0; Mismatches 0; Indels 54; Gaps 2;
1 MILEERPDGAGAGESPRLOISRRKPRKTRVSSLRGRREGRLDVCAMGRCVHRGEDPVR 60
1 MILEERPDGAGAGESPRLO----- 20
61 VHVGPMPOLHAGVCDLSLTQCGLOSRRAQIHQOIKDELQMTGAEINLYRATSNRVR 120
21 -----GCDLSLTQCGLOSRRAQIHQOIKDELQMTGAEINLYRATSNRVR 67
121 ETVALELSYVNSNLQKLEELSGVDPGRHSGEAVTVPMIPLGLKETKELDWSTPLK 180
68 ETVALELSYVNSNLQKLEELSGVDPGRHSGEAVTVPMIPLGLKETKELDWSTPLK 127
181 ELISVHFGEDGASVFAEIRELEALRQAMRTPSRNESGLELLTAYYNQICFLDARFLTPAR 240
128 ELISVHFGEDGASVFAEIRELEALRQAMRTPSRNESGLELLTAYYNQICFLDARFLTPAR 187

QY 241 SLGLFFHYDYSITGVPAQORALAFKSGVLEFNIGALHTQIGARODRSCCTEGARRAMEAFQ 300
 DB 188 SLGLFFHYDYSITGVPAQORALAFKSGVLEFNIGALHTQIGARODRSCCTEGARRAMEAFQ 247
 QY 301 RAAAGAFSLRNFHSHAPSPDMSAASLCALQELMMAQAQECVFEGLSPPASMAPODCLAQL 360
 DB 248 RAAAGAFSLRNFHSHAPSPDMSAASLCALQELMMAQAQECVFEGLSPPASMAPODCLAQL 307
 QY 361 RLAEAAQVAAEYRLVHRTMAQPPVHDYVPSWTALVHVKAIEYFRSLAHYHVAMALCDGS 420
 DB 308 RLAEAAQVAAEYRLVHRTMAQPPVHDYVPSWTALVHVKAIEYFRSLAHYHVAMALCDGS 367
 QY 421 PATEGELPHEQVFLQPPPTSSKPRGVPVLPQELERRRQLGKAHLKRAILGQEEALRLHALC 480
 DB 368 P-TEGELPHEQVFLQPPPTSSKPRGVPVLPQELERRRQLGKAHLKRAILGQEEALRLHALC 426
 QY 481 RVLEVDLLRAVISQTLQSLAKYAELEDDDFCEAAEAPDIQPKTHQKPEARMPLRSQ 540
 DB 427 RVLEVDLLRAVISQTLQSLAKYAELEDDDFCEAAEAPDIQPKTHQKPEARMPLRSQ 486
 QY 541 KGPDIFHRLGFLSVFSAKNRWLVGPVHLTGEGGFGTLRGDSPVLIAAVIPGSOAAAA 600
 DB 487 KGPDIFHRLGFLSVFSAKNRWLVGPVHLTGEGGFGTLRGDSPVLIAAVIPGSOAAAA 546
 QY 601 GLKEGDYIVSVNGQPCRWHRHAEVVTTELKAAAGEAGASLQVVSLLPSSRLSLGDRRPVLL 660
 DB 547 GLKEGDYIVSVNGQPCRWHRHAEVVTTELKAAAGEAGASLQVVSLLPSSRLSLGDRRPVLL 606
 QY 661 GPRGLLSQREHGCKTASTWASPRLLNWSRKAQOQKTCGCPQCAPVPKPPSSLKHP 720
 DB 607 GPRGLLSQREHGCKTASTWASPRLLNWSRKAQOQKTCGCPQCAPVPKPPSSLKHP 666
 QY 721 GWP 723
 DB 667 GWP 669

RESULT 5

US-10-697-266-4
 ; Sequence 4, Application US/10697266
 ; Publication No. US20040067522A1
 ; GENERAL INFORMATION:
 ; APPLICANT: RUSCH, Douglas et al
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; FILE REFERENCE: CL001099-CIP-DIV2
 ; CURRENT APPLICATION NUMBER: US/10/697,266
 ; PRIOR FILING DATE: 2003-10-31
 ; PRIOR APPLICATION NUMBER: 10/274,878
 ; PRIOR FILING DATE: 2002-10-22
 ; PRIOR APPLICATION NUMBER: 09/849,834
 ; PRIOR FILING DATE: 2001-05-07
 ; PRIOR APPLICATION NUMBER: 09/773,371
 ; PRIOR FILING DATE: 2001-02-01
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 634
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-10-697-266-4

Query Match 61.8%; Score 2331; DB 12; Length 634;
 Best Local Similarity 67.0%; Pred. No. 2,7e-195;
 Matches 473; Conservative 49; Mismatches 112; Indels 72; Gaps 6;
 QY 1 MILERDPGAGAGESPRLQISRKPRKTRVSSLRGREGRLDVCWARGCRVHGEDPVR 60
 DB 1 MILERDPGAGAGESPRLQISRKPRKTRVSSLRGREGRLDVCWARGCRVHGEDPVR 60
 QY 61 VHVGPMPQLHVGCDLSITQCGQLQSRRAHQHIDKELQMTGAENLYRATSNRVR 120

DB 27 -----KGYGSFVNQPCQLQSHRRLHQQLSKELRMETGAENLYRATSNRVR 74
 QY 121 ETVALELSYNSNLOLLKEELELSGVDVDPGRHGESEAVTVMIPGLKTKELDWSTPLK 180
 DB 75 ETVALELSYNSNLOLLKEELELSGVDVDPGRHGESEAVTVMIPGLKTKELDWSTPLK 134
 QY 181 ELISVHGEDGASVEAIRLEALRQAMRTSPSRNESGLELLTAYYNOLCFDARFLPAR 240
 DB 135 ELISEHGEDGTSEFEIQELEDLRQATRTSPSRNESGLELLTAYYNOLCFDARFLPAR 194
 QY 241 SLGLFFHYDYSITGVPAQORALAFKSGVLEFNIGALHTQIGARODRSCCTEGARRAMEAFQ 300
 DB 195 SPGLFFHYDYSITGVPAQORALAFKSGVLEFNIGALHTQIGARODRSCCTEGARRAMEAFQ 254
 QY 301 RAAAGAFSLRNFHSHAPSPDMSAASLCALQELMMAQAQECVFEGLSPPASMAPODCLAQL 360
 DB 255 RAAAGAFSLRNFHSHAPSPDMSAASLCALQELMMAQAQECVFEGLSPPASMAPODCLAQL 314
 QY 361 RLAEAAQVAAEYRLVHRTMAQPPVHDYVPSWTALVHVKAIEYFRSLAHYHVAMALCDGS 420
 DB 315 RLAEAAQVAAEYRLVHRTMAQPPVHDYVPSWTALVHVKAIEYFRSLAHYHVAMALCDGS 374
 QY 421 PATEGELPHEQVFLQPPPTSSKPRGVPVLPQELERRRQLGKAHLKRAILGQEEALRLHALC 480
 DB 375 PA-KGELARQEHVF-QPSTPHEPLGPTLPQHPEDRRKLAKAHLKRAILGQEEALRLHALC 432
 QY 481 RVLEVDLLRAVISQTLQSLAKYAELEDDDFCEAAEAPDIQPKTHQKPEARMPLRSQ 540
 DB 433 RVLEVDLLRAVISQTLQSLAKYAELEDDDFCEAAEAPDIQPKTHQKPEARMPLRSQ 483
 QY 541 KGPDIFHRLGFLSVFSAKNRWLVGPVHLTGEGGFGTLRGDSPVLIAAVIPGSOAAAA 600
 DB 484 -----GPLSVFSTKNRWLVGPVHMTGEGGFGTLRGDSPVLIAAVIPGSOAAAA 534
 QY 601 GLKEGDYIVSVNGQPCRWHRHAEVVTTELKAAAGEAGASLQVVSLLPSSRLSLGDRRPVLL 660
 DB 535 GLKEGDYIVSVNGQPCRWHRHAEVVTTELKAAAGEAGASLQVVSLLPSSRLSLGDRRPVLL 594
 QY 661 GPRGLLSQREHGCKTASTWASPRLLNWSRKAQOQKTCGCPQCAPVPKPPSSLKHP 706
 DB 595 W-----NORECGFETPMPTRTRPWPILGWSRKNKQKGTGSHHPDC 634

RESULT 6

US-10-274-878-4
 ; Sequence 4, Application US/10274878
 ; Publication No. US20030049792A1
 ; GENERAL INFORMATION:
 ; APPLICANT: RUSCH, Douglas et al
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; FILE REFERENCE: CL001099-CIP-DIV
 ; CURRENT APPLICATION NUMBER: US/10/274,878
 ; PRIOR FILING DATE: 2002-10-22
 ; PRIOR APPLICATION NUMBER: 09/849,334
 ; PRIOR FILING DATE: 2001-05-07
 ; PRIOR APPLICATION NUMBER: 09/773,371
 ; PRIOR FILING DATE: 2001-02-01
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 634
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-10-274-878-4

Query Match 61.8%; Score 2331; DB 14; Length 634;
 Best Local Similarity 67.0%; Pred. No. 2,7e-195;
 Matches 473; Conservative 49; Mismatches 112; Indels 72; Gaps 6;
 QY 1 MILERDPGAGAGESPRLQISRKPRKTRVSSLRGREGRLDVCWARGCRVHGEDPVR 60

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2004, 11:14:21 ; Search time 83 Seconds
(without alignments)
2748.430 Million cell updates/sec

Title: US-10-697-266-2
Perfect score: 3774
Sequence: 1 MILEERPDGAGAESPRLO.....QPCAPVKPAPPSLKHPGWP 723

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SPTEMBL_25:*
- 2: sp_archaea:*
- 3: sp_bacteria:*
- 4: sp_fungi:*
- 5: sp_human:*
- 6: sp_invertebrate:*
- 7: sp_mammal:*
- 8: sp_mhc:*
- 9: sp_organelle:*
- 10: sp_plant:*
- 11: sp_prodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3449.5	91.4	670	Q8TAV1	Q8tav1 homo sapien
2	1847	48.9	410	Q96PV9	Q96pv9 homo sapien
3	1496	39.6	451	Q80WU2	Q80wu2 mus musculus
4	1316	34.9	683	Q803B1	Q803b1 brachydanio
5	1284	34.0	686	Q8BXG3	Q8bxg3 canis famil
6	1255	33.3	686	Q8NE33	Q8ne33 homo sapien
7	1255	33.3	686	Q8IUC4	Q8iuc4 homo sapien
8	1242	32.9	686	Q9DEN2	Q9dbn2 mus musculus
9	1237	32.8	657	Q8N3T7	Q8n3t7 homo sapien
10	1228	32.5	686	Q8BWR8	Q8bw-8 mus musculus
11	1185.5	31.4	685	Q9GRU1	Q9gru1 homo sapien
12	1023	27.1	718	Q9XYU9	Q9xyu9 drosophila
13	887.5	23.5	687	Q9VXP2	Q9vxp2 drosophila
14	394	10.4	298	Q8N9D6	Q8n9d6 homo sapien
15	387	10.3	1421	Q9P257	Q9p257 homo sapien
16	380.5	10.1	1636	Q9H3S7	Q9h3s7 homo sapien

17	335.5	8.9	794	5	Q8T7K0	Q8t7k0 dictyosteli
18	333	8.8	846	10	Q8H1H8	Q8h1h8 arabidopsis
19	295.5	7.8	883	10	Q8LNU2	Q8lnu2 oryza sativ
20	292.5	7.8	773	3	Q8WZL4	Q8wzl4 yarrowia li
21	289	7.7	816	10	Q9X156	Q9xi56 arabidopsis
22	285.5	7.6	775	3	Q13783	Q13783 schizosacch
23	284.5	7.5	846	5	Q8I114	Q8i114 caenorhabdi
24	284.5	7.5	882	5	Q8I115	Q8i115 caenorhabdi
25	284	7.5	465	13	Q7SY03	Q7sy03 brachydanio
26	276.5	7.3	498	5	Q7YXC7	Q7yxc7 caenorhabdi
27	266.5	7.1	872	11	Q8Y09	Q8y09 mus musculu
28	262.5	7.0	1833	5	Q9VUH6	Q9vuh6 drosophila
29	262.5	7.0	1838	5	Q960G3	Q960g3 drosophila
30	255.5	6.8	847	3	P79020	P79020 emericeila
31	239	6.3	836	5	Q9VB05	Q9vb05 drosophila
32	232	6.1	661	3	Q12033	Q12033 saccharomyc
33	224.5	5.9	266	5	Q9NA90	Q9na90 caenorhabdi
34	203.5	5.4	785	3	Q9UW12	Q9uw12 candida alb
35	165.5	4.4	1802	5	Q9VQU8	Q9vqu8 drosophila
36	165	4.4	1190	5	Q9V562	Q9v562 drosophila
37	165	4.4	1205	5	Q8MKS8	Q8mks8 drosophila
38	162	4.3	1386	5	Q8MKS7	Q8mks7 drosophila
39	162	4.3	1407	5	Q8IMD0	Q8imd0 drosophila
40	155.5	4.1	942	4	Q8IUV5	Q8iuv5 homo sapien
41	155.5	4.1	3753	2	Q846W6	Q846w6 streptomyce
42	151	4.0	1024	11	Q9ESJ5	Q9esj5 mus musculu
43	149	3.9	799	13	Q7SYH9	Q7syh9 gallus gall
44	147	3.9	1151	2	Q84EM0	Q84em0 rhodobacter
45	144	3.8	966	4	Q8NDG1	Q8ndg1 homo sapien

ALIGNMENTS

RESULT 1

ID	Q8TAV1	PRELIMINARY;	PRT;	670 AA.
AC	Q8TAV1;			
DT	01-JUN-2002 (Tremblrel. 21, Created)			
DT	01-JUN-2002 (Tremblrel. 21, Last sequence update)			
DT	01-JUN-2003 (Tremblrel. 24, Last annotation update)			
DE	Similar to GTP-rho binding protein 1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Theria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Blood;			
RA	Strausberg R.;			
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.			
CC	-!- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.			
DR	EMBL; BC025767; AAH25767.1; -			
DR	GO; GO:0005622; C:intracellular; IEA.			
DR	GO; GO:0007242; P:intracellular signaling cascade; IEA.			
DR	InterPro; IPR004328; BR01.			
DR	InterPro; IPR001478; PDZ.			
DR	InterPro; IPR00861; REM_repeat.			
DR	Pfam; PF03097; BR01; 1.			
DR	Pfam; PF02185; HR1; 1.			
DR	Pfam; PF00595; PDZ; 1.			
DR	SMART; SM00074; HR1; 1.			
DR	SMART; SM00228; PDZ; 1.			
DR	PROSITE; PS50106; PDZ; 1.			
SQ	SEQUENCE 670 AA; 73589 MW; CE8C01B716465FF0 CRC64;			

Query Match 91.4%; Score 3449.5; DB 4; Length 670;
Best Local Similarity 92.7%; Pred. No. 1.2e-239;
Matches 670; Conservative 0; Mismatches 0; Indels 53; Gaps 1;

Oy 1 MILEERPDGAGAESPRLOISRRKPKTRVSSLRGRREGLRDVCANRGCRVHGDFVR 60
Db 1 MILEERPDGAGAESPRLO-----

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC052010; AAH52010.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR004328; BRO1.
DR InterPro; IPR000861; REM_repeat.
DR Pfam; PF03097; BRO1; 1.
DR Pfam; PF02185; HRI; 1.
DR SMART; SM00074; HRI; 1.
DR SMART; SM00228; PDZ; 1.
DR PROSITE; PS0106; PDZ; 1.
SQ SEQUENCE 451 AA; 49072 MW; D69B6A232CA19018 CRC64;

Query Match 39.6%; Score 1496; DB 11; Length 451;
Best Local Similarity 67.3%; Pred. No. 2.7e-99;
Matches 309; Conservative 28; Mismatches 74; Indels 48; Gaps 4;

QY 1 MILERRPDGAGGEEPRIOISRRKPKTRVSLRGREGLRDVCAWRCGRVHRGDPVR 60
DB 1 MILERRPDGQGGERSRRPQDD-----GSIR----- 26

QY 61 VHVGPMPQLHVGCDLSLTQICGQLOSRRAHQHIDKELQWRTGAENLYRATSNRYR 120
DB 27 -----KGYGSFVQNPQGGQSHRARLHQHISKELRWRTGAENLYRATSNRYR 74

QY 121 ETVALELSVYNSNLOLLKEELSGVDPGRHGSFAVTPMPLGLKTKELDWSPLK 180
DB 75 ETVALELSVYNSNLOLLKEELSLSTVDVDQPEGGITIPMPLGLKTKELDWAIPLK 134

QY 181 ELISVHFGEDGASYEAEIRELEALRQAMTSPRNFSGLELLTAYYNOLCFDARFLTPAR 240
DB 135 ELISVHFGEDGTSFETIELEDLAQAATPTPSRDEAGDLLAAYISQLCFDARFSPSR 194

QY 241 SGLFPHWYDLSLTGVPQAORALAFKGVSLFNIGALHTQIGARQDRSCTEGARAMEAFQ 300
DB 195 SPGLLFHWYDLSLTGVPQAORALAFKGVSLFNIGALHTQIGARQDRSCTEGTNHAARAFQ 254

QY 301 RAAGAFSLLENFNSHAPSPDMSAASLCALQOLMAQAQSCVFEGLSPSPASMAPQCLAQ 360
DB 255 RAAGAFSLLENFNSHAPSPDMSAASLSMLEQNTAQACIECFKGLLPASATPDCPDQL 314

QY 361 RLAEAAQAAVAAEYRLVHRTMAQPPVHDYVPVSWTALVHVKAHYFRSLAHYHVMALCDGS 420
DB 315 QLAQAAQAAVAAEYGLVHRAMAQPPVVDYLPASWTNLAVKAEHFCALAHYHVMALCESH 374

QY 421 PATEGELTHQVFLQPTTSSKPRGPVLPOELEERRQLG 459
DB 375 PA-KGELARQEHVF-QPSTPHEPLGFTLQHPEDRRKLG 411

RESULT 4

Q803B1
ID Q803B1 PRELIMINARY; PRT; 683 AA.
AC Q803B1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to rhophilin-2.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC044556; AAH44556.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR004328; BRO1.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR000861; REM_repeat.
DR Pfam; PF03097; BRO1; 1.
DR Pfam; PF02185; HRI; 1.
DR Pfam; PF00595; PDZ; 1.
DR SMART; SM00074; HRI; 1.
DR SMART; SM00228; PDZ; 1.
DR PROSITE; PS0106; PDZ; 1.
SQ SEQUENCE 683 AA; 77001 MW; B06F4D2B088FCB52 CRC64;

Query Match 34.9%; Score 1316; DB 13; Length 683;

Best Local Similarity 44.3%; Pred. No. 4.5e-86;

Matches 271; Conservative 122; Mismatches 205; Indels 14; Gaps 4;

QY 41 LRDVCAWRCGRVHRGDPVRVHVGPMPQLHVGCDLSLTQICGQLOSRRAHQHIDKE 100
DB 1 MTDLTLPNGCKSN-----GPVENVGYFKGCNPNFAQTGSKLONKRAVLNOKLIQ 50

QY 101 LQWRTGAENLYRATSNRYRRTVALELSVYNSNLOLLKEELSGVDPGRHGSFAVTP 160
DB 51 MMRRAGAENLLKATSNKRVREQVLELSYVNSDLQLLMEQLGELNSVVEYQNVQSSSI 110

QY 161 PMPLGLKTKELDWSPLKELISVHFGEDGASYEAEIRELEALRQAMTSPRNFSGLEL 220
DB 111 PLPLGLKTKELDWSVPLKDFILEHSEDSGNFQNDLMDLQACQTPSRNNSGVDL 170

QY 221 LPAYYNOLCFDARFLTPARSLGLFHHWYDLSLTGVPQAORALAFKGVSLFNIGALHTQI 280
DB 171 IANYFSQLSFLTRFFSPTRQIGIFFTWDSFTGMPVCQNNISLEKASMLFNMAALYSQI 230

QY 281 GARQDRSCTEGARAMEAFQRAAGAFSLLENFNSHAPSPDMSAASLCALQOLMAQAQSC 340
DB 231 GTRADRQTLAGLEDAIAAFQKSAAGVHLHLLKETFTHTPTPSYDMSFAMLSMLFRMLAQASC 290

QY 341 VPEGLSPSPASMAPQCLAQRLAQAAQVAAEYRLVHRTMAQPPVHDYVPVSWTALVHVK 400
DB 291 LFEQITLPGIRNEFFCL--LMAQAAKAAVAAEYAOVHOSMIETPTKNVFFWTTMSQLK 348

QY 401 AEYFRSLAHYHVMALCDGSPATEGELTHQVFLQ-PFTSSKPRGPV-LPOELEERRQL 458
DB 349 INHYNLSAHYFVSTALLDHLQNPSSDEKQKALSQLYDAMPEGSPDLILKNKERRRI 408

QY 459 GKALHRAILQGERALRLHALCRVLREVDLRAVLSQTLQSLAKYAELEDREDDFCEAAE 518
DB 409 GKALQRSIMGHEAIRTCHRRRLQKLDILSLIRASLNRSLTKFEQNDKEDFTDYL 468

QY 519 APDIQPKTHQKPEARMPLRSQKGPDIHRLGPLSVFSKXNRWRLVGPVHLTRGEGGCL 578
DB 469 APDIISKTEKAEIMEIPATKVKVTDLQRLGPLSVFSKQRTWAPRTIRILQDRDLGF 528

QY 579 TLRGDSPVLIAAVTPGSOAAAAGLKEGDYIVSVNGQPCQRRWRHAEVVTTELKAAAGAGASL 638
DB 529 TLKGDAPVQIQSLDPLCPAAAGGLKEGDYLVAVGDTCKWNGVSDVMKLLKDVDERGINI 588

QY 639 QVWSLLPSSRLP 650
DB 589 RVVSMMDSSSQP 600

RESULT 5

Q8HXG3

ID Q8HXG3 PRELIMINARY; PRT; 686 AA.

Q8HXG3; 559 VSIQDVDCWLTVTSEVMKGLKACGRDGVEMKVVSL 594
01-MAR-2003 (Tremblrel. 23, Created)
01-MAR-2003 (Tremblrel. 23, Last sequence update)
01-JUN-2003 (Tremblrel. 24, Last annotation update)
P76RBE protein.
Canis familiaris (Dog).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
NCBI_TaxID=9615;
[1]
SEQUENCE FROM N.A.
MDLINE=22361238; PubMed=12473120;
Mircescu H., Steuve S., Savonet V., Degraef C., Mellor H.,
Dumont J.E., Maenhaut C., Pison I.;
"Identification and characterization of a novel activated Phob binding
protein containing a PDZ domain whose expression is specifically
modulated in thyroid cells by cAMP.";
Eur. J. Biochem. 269:6241-6249 (2002).
EURL; AJ347749; CAC87938.1; -.
EURL; AJ347749; CAC87938.1; -.
GO; GO:0005622; C:intracellular; IEA.
GO; GO:0007242; P:intracellular signaling cascade; IEA.
InterPro; IPR004328; BRO1.
InterPro; IPR001478; PDZ.
InterPro; IPR000861; REM_repeat.
Pfam; PF03097; BRO1; 1.
Pfam; PF02185; HR1; 1.
Pfam; PF00595; PDZ; 1.
SMART; SM00074; HR1; 1.
SMART; SM00228; PDZ; 1.
PROSITE; PS0106; PDZ; 1.
SEQUENCE 686 AA; 76965 MW; DA554284E53C4318 CRC64;
Query Match 34.0%; Score 1284; DB 4; Length 686;
Best Local Similarity 47.6%; Pred. No. 9.1e-84;
Matches 274; Conservative 96; Mismatches 196; Indels 10; Gaps 4;
QY 74 GDSLTQICQGLQSRRAQIHOQIDKELQMTGAENLYRATSNRRVETVALELSVNSN 133
DB 24 GCNPLAQTRSKLQNRALNQILKAVRMRTGAENLLKVAINHVKVRQVRLSESVNSD 83
QY 134 LQLKEELEELSGGVDPRGHSGEAVTPMIPGLKTKELDWSTPLKELISVHFEDGAS 193
DB 84 LQMLKEELEGLNISGVYQSTEEAFTPLIPGLKTKEDIDFSVVLKDFILEHSEDSYL 143
QY 194 YEAEITRELEALROAMRTSPNESGLELLTAYYNQICFLDARFLTPARSIGLFFHHYDSLT 253
DB 144 YEDEIADLMDLRQACRTSPDEAGVELLMTYFIQLGFVESRFPFPTRQMGILLFTWYDSLT 203
QY 254 GYPAQORALAFKGSVLFNIGALHTQIGARODRSCTEGARRAMEAFORAAAGAFSLLENF 313
DB 204 GYVPSQONLLLEKASVLFNTGALYTIQIGTRCNRTOAGLDGAVDAFQRAAGVLLHLKETP 263
QY 314 SHAPSPDMSAALCALQELMMAQAQECVFEGLSPASMAPQCLQLRLAQAQAQVAAEY 373
DB 264 THTPSYDMSAPMLSVLVKMLAQAEQNFVKLPFG--IRNEFFVLVVKVAQAQAQVGEVY 321
QY 374 RLVRHTMAQPPVHDVVPVSWTALMVHVKARYPSLAHYHVMALCDG--SPATEGELPTHE 431
DB 322 RLHTAMSQAPVKENIPYSWASLVCKVAHYAALAHYFAATLLIDHQLKPGADED---HQ 378
QY 432 QVFLQPTSSKPRG--PVLP--QELERRQLGKAHLKRAILGOEALRLHALCRVLEVDL 488
DB 379 EKCLSQLYDHMPPEGLTFLATLKSQHQRQLGKSHLRRAVAHHESVREASLCKKLRIEY 438
QY 489 LRAVTSQTLQSLAKYAEALDREDDFCEAAEAPDIQKTHQKPEARMPLRSQCKGPDIFHR 548
DB 439 LQDVLVSVAERSRLKYAQHQDDDLNLNLIDAPDIIISKTEQEVETIILPQFSKVTATDFQK 498
QY 549 LGPLSVFSANKRWLVGPHLRTGEGGFLTLRGDSFVLIAAIVIPGSAQAAGLKEGDYI 608
DB 499 LGPLSVFSANKRWTPPRSIHFTHAEEDGLGFTLRGNSPVQVHFLDPHCSAALAGKEGDYI 558
QY 609 VSYNGQFCRWRHAEVVTTELKAAAGEAGASLQVVSL 644

Q8NE33; 559 VSIQDVDCWLTVTSEVMKGLKACGRDGVEMKVVSL 594
01-OCT-2002 (Tremblrel. 22, Created)
01-OCT-2002 (Tremblrel. 22, Last sequence update)
01-JUN-2003 (Tremblrel. 24, Last annotation update)
Similar to rhophilin-like protein.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=Testis;
Strausberg R.;
Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
EMBL; BC036447; AAH36447.1; -.
GO; GO:0005622; C:intracellular; IEA.
GO; GO:0007242; P:intracellular signaling cascade; IEA.
InterPro; IPR004328; BRO1.
InterPro; IPR001478; PDZ.
InterPro; IPR000861; REM_repeat.
Pfam; PF03097; BRO1; 1.
Pfam; PF02185; HR1; 1.
SMART; SM00074; HR1; 1.
SMART; SM00228; PDZ; 1.
SEQUENCE 686 AA; 76992 MW; F13AABBF958947F5 CRC64;
Query Match 33.3%; Score 1255; DB 4; Length 686;
Best Local Similarity 45.9%; Pred. No. 1.1e-81;
Matches 266; Conservative 107; Mismatches 196; Indels 10; Gaps 5;
QY 74 GDSLTQICQGLQSRRAQIHOQIDKELQMTGAENLYRATSNRRVETVALELSVNSN 133
DB 24 GCNPLAQTRSKLQNRALNQILKAVRMRTGAENLLKVAINHVKVRQVRLSESVNSD 83
QY 134 LQLKEELEELSGGVDPRGHSGEAVTPMIPGLKTKELDWSTPLKELISVHFEDGAS 193
DB 84 LQMLKEELEGLNISGVYQSTEEAFTPLIPGLKTKEDIDFSVVLKDFILEHSEDSYL 143
QY 194 YEAEITRELEALROAMRTSPNESGLELLTAYYNQICFLDARFLTPARSIGLFFHHYDSLT 253
DB 144 YEDEIADLMDLRQACRTSPDEAGVELLMTYFIQLGFVESRFPFPTRQMGILLFTWYDSLT 203
QY 254 GYPAQORALAFKGSVLFNIGALHTQIGARODRSCTEGARRAMEAFORAAAGAFSLLENF 313
DB 204 GYVPSQONLLLEKASVLFNTGALYTIQIGTRCNRTOAGLDGAVDAFQRAAGVLLHLKETP 263
QY 314 SHAPSPDMSAALCALQELMMAQAQECVFEGLSPASMAPQCLQLRLAQAQAQVAAEY 373
DB 264 THTPSYDMSAPMLSVLVKMLAQAEQNFVKLPFG--IRNEFFVLVVKVAQAQAQVGEVY 321
QY 374 RLVRHTMAQPPVHDVVPVSWTALMVHVKARYPSLAHYHVMALCDG--SPATEGELPTHE 431
DB 322 QQLHAAMSQAPVKENIPYSWASLVCKVAHYAALAHYFAATLLIDHQLKPGADED---HQ 378
QY 432 QVFLQPTSSKPRG--PVLP--QELERRQLGKAHLKRAILGOEALRLHALCRVLEVDL 488
DB 379 EKCLSQLYDHMPPEGLTFLATLKSQHQRQLGKSHLRRAVAHHESVREASLCKKLRIEY 438
QY 489 LRAVTSQTLQSLAKYAEALDREDDFCEAAEAPDIQKTHQKPEARMPLRSQCKGPDIFHR 548
DB 439 LQDVLVSVAERSRLKYAQHQDDDLNLNLIDAPDIIISKTEQEVETIILPQFSKVTATDFQK 498
QY 549 LGPLSVFSANKRWLVGPHLRTGEGGFLTLRGDSFVLIAAIVIPGSAQAAGLKEGDYI 608
DB 499 LGPLSVFSANKRWTPPRSIHFTHAEEDGLGFTLRGNSPVQVHFLDPHCSAALAGKEGDYI 558
QY 609 VSYNGQFCRWRHAEVVTTELKAAAGEAGASLQVVSL 644


```
QY 194 YEAREIRELEALROAWRTSPRESNEGLELLTAYYNOLCFDARFLTPARSLGFFHWDYSLT 253
DB 144 YEDDIAUDMDURQACRTSPRDEAGVELLMSTFIQLGFEVSFFPPTHMGLLFTWYDSFT 203
QY 254 GVPAAQORALAPEKGSVFNIGALHTQIGARQDRCTEGARRAMEAFQRAAGAFSLLENF 313
DB 204 GVPVSQOTLLLEKASVLFNIGALYTIQIGTRCNROTQAGLESADVAFQRAAGVNLVKETP 263
QY 314 SHAPSPMSAASLALCALQOLMAQAQECVFEGLSPASMAPODCLAQLAEAAQVAEY 373
DB 264 THTESYDMSPAVLSVLMKMLAQAESVFEKVLPG--IQNEFFVLMVKAQEAQVAEY 321
QY 374 RLVRHTMAQPPVHDVVPVSWTALVHVKAIEYFSLAHYHVAMALCDG--SPATEGELPTHE 431
DB 322 RQLHAAMSQEPVENIPYSWASVAVKAYHYGALAHYFAATLLIDHQLKPGADE--HQ 378
QY 432 QVFLQPPTSSKPRGVLVPELE--ERRQLGKAHLKRAILGOEALRHLCRVLRVDL 488
DB 379 EKCLSQLYDRMEGWTPLATLKNAGQVLLGKHLHRAIGFHEESLEBANLCKKLRIQV 438
QY 489 LRAVISQTLQRLSLAKYAELEDREDDCEAAABADIPQKTHQKPEARMPLRSQKGPDIHR 548
DB 439 LEDVLSAAHQRTQLKHTQHRREDLLNLIDAPDLPKTEREVKIITPFDKVTVDFFQK 498
QY 549 LGPLSVFSKKNRWLRVGVPHLTRGEGFGLTLRGDSPVLIAAVIPGSQAAAAGLKEGDYI 608
DB 499 LGPLSVFSKKNRWSPRGHIFVTEGDLGFTLRGNTPVQVHFLDPHCSASLAGAKEDYI 558
QY 609 VSVNGQPCRWHRHAEVVTTELKAAAGASQVVSLLPSS 647
DB 559 VSIQGVDCWKLTVSEVMKILKSGFGEVEMKVVSLLDST 597

RESULT 9
Q8N3T7 PRELIMINARY; PRT; 657 AA.
AC Q8N3T7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKFZP761L1918.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Amygdala;
RA Ottenwaelder B., Obermaier B., Mewes H.W., Weil B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL831950; CAD38597.1;
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR004328; BRO1.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR000861; REM_repeat.
DR Pfam; PF03097; BRO1; 1.
DR Pfam; PF02185; HRI; 1.
DR SMART; SM00074; HRI; 1.
DR SMART; SM00228; PDZ; 1.
DR KFT Hypothetical protein.
RW NON_TER
SQ SEQUENCE 657 AA; 73850 MW; 562DEF16F238B39D CRC64;

Query Match 32.8%; Score 1237; DB 4; Length 657;
Best Local Similarity 46.1%; Pred. No. 2.1e-80;
Matches 262; Conservative 106; Mismatches 190; Indels 10; Gaps 5;

QY 85 QLOSRAQIHQOQDKELQMTGAGNLYRATSNRNKRVETVALELSYVNSNLQKLKEELEL 144
DB 6 KLQNGRAALNCQILKAVRMRTGAENLLKVATNSKRVQVRLLELSFVNSDLQMLKEELEGL 65
```

```
QY 145 SGGVDPGRHSGEAVTVPMIPGLKTKETKELDWSTPLKELISVHFEGDGSYAEIRELEAL 204
DB 66 NISGVYQNTTEAFTIPLIPLGXETKDVDFAVVVKDFILEHYSEDGYLXDEIADLMDL 125
QY 205 ROAMRTSPRESNEGLELLTAYYNOLCFDARFLTPARSLGFFHWDYSLTGVPAQORALAF 264
DB 126 RQAKRTSPRDEAGVELLMSTFIQLGFEVSFFPPTHMGLLFTWYDSLTGVVSQONLL 185
QY 265 EKGSVLNFIGNALHTQIGARQDRCTEGARRAMEAFQRAAGAFSLLENFSHAPSPDMSAA 324
DB 186 EKASVLFNTGALYTIQIGTRCDROTQAGLESADAFQRAAGVNLKDTFTHTPSYDMSPA 245
QY 325 SLCALQOLMAQAQECVFEGLSPASMAPODCLAQLAEAAQVAEYRLVHRTMAQPP 384
DB 246 MLSVLVKKMLAQAESVFEKISLPG--IRNEFFVLMVKAQEAQVGEYQQLHAAMSQAP 303
QY 385 VHDYVPVSWTALVHVKAIEYFSLAHYHVAMALCDG--SPATEGELPTHEQVFLQPPTSSK 442
DB 304 VKENIPYSWASLACVKAHYAALAHYFAATLLIDHQLKPGTDLD--HQECLSLQLYDHM 360
QY 443 PRG--PVLP--QELERRQLGKAHLKRAILGOEALRHLCRVLRVDLRLRAVISQTLQ 499
DB 361 PEGLTPLATLKNQDQRRQLGKSHLRAMAHHEESVREASLCKLRSIEVLQKVLCAQBR 420
QY 500 SLAKYAELEDREDDCEAAABADIPQKTHQKPEARMPLRSQKGPDIHRLGPLSVFSAKN 559
DB 421 SRLTYAQHEEDLLNLIDAPSVVAKTEQVDIILPQFSKLTVTDFQKLGPLSVFSANK 480
QY 560 RWLRVGVPHLTRGEGFGLTLRGDSPVLIAAVIPGSQAAAAGLKEGDYIVSVNGQPCRW 619
DB 481 RWTTPRSRFTAEAGDLGFTLRGNAPVQVHFLDPYCSASVAGAREGDYIVSIQLVDCWKL 540
QY 620 RHAEVVTTELKAAAGASQVVSLLPSS 647
DB 541 TLSEVMKILKSGFGEVEMKVVSLLDST 568

RESULT 10
Q8BWR8 PRELIMINARY; PRT; 686 AA.
AC Q8BWR8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Similar to rhophilin-like protein.
GN RHPN2 OR D7ERTD784E.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002).
RL Nature 420:563-573 (2002).
RL EMBL; AK050214; BAC34127.1;
RL MGD; MGI:1289234; Rhp2.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR004328; BRO1.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR000861; REM_repeat.
DR Pfam; PF03097; BRO1; 1.
DR Pfam; PF02185; HRI; 1.
DR Pfam; PF00595; PDZ; 1.
DR SMART; SM00074; HRI; 1.
DR SMART; SM00228; PDZ; 1.
DR PROSITE; PS50106; PDZ; 1.
SQ SEQUENCE 686 AA; 76905 MW; 49B67AA361FEDC58 CRC64;
```



```
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -| SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.
DR EMBL; AE003500; AAF48516.2; -.
DR FlyBase; FBgn0026374; Rhp.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR004328; BR01.
DR InterPro; IPR001478; PDZ.
DR Pfam; PF03097; BRO1; 1.
DR Pfam; PF00595; PDZ; 1.
DR SMART; SM00228; PDZ; 1.
DR PROSITE; PS0106; PDZ; 1.
SQ SEQUENCE 687 AA; 77343 MW; C83893F0C8C9E6D2 CRC64;

Query Match      23.5%; Score 887.5; DB 5; Length 687;
Best Local Similarity 31.8%; Pred. No. 3.2e-55;
Matches 228; Conservative 120; Mismatches 228; Indels 141; Gaps 19;

QY 54 RGDPRVRVHVPMPOLHVGCDLSLTQICGQLQSRRAQIHQOIKELQWRTGAENLYRA 113
Db ||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
33 RGSDFP-----RAATCS-----KLNRRCKLNKEINKELRURAGAENLYK- 72
QY 114 TSNRRVRETVALELSYVNSLNQLKEELSGVDPGRHSGEAVTVPMIPLGLKETKEL 173
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
73 -----LNSSVEIYQSE-----SHNGIMPMPGLKETKEI 102

QY 174 DNSTPLKELISVHFGDGSAYEIRELEALQAMTPSPNESGLELLTAYYNQLCFDA 233
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
103 NFMEPPSDFILEHSEPSMYIDAIDMTDTROASKTPSRDALGVALLFRYYNTLYYVER 162

QY 234 RFLTPARSLGLPFWYDLSLTGVPQAORALAFKGSVLNIGALHTOIGARQDRSCTEGAR 293
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
163 RFPFPDENLGVFEWYDLSLTGVSCQRTAFKACTILNFGGYYTQIGARHDTTTERGLD 222

QY 294 RAMEAFQRAAGAFSLLENFNSHAPSDMSAASICALQLMAQAQECVPEGLSPASMAP 353
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
223 LAVDSFLRAAGVFRHYDTFTNAPSDMLKPVQDLVLSMLSQARECLFEKLOIQTEAMS 282

QY 354 QCLAQRLAQEAQAAVEYRLVHRTWAQPPHYDVPVSWTALVHVHKAQYFRLAHYHVA 413
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
283 HDCQAFRDLAGEAAQISHEYNEMHONIQANDHTYLPCEWAGLVVPVKAELYKAFAHFYKA 342

QY 414 MALCDGSPATEGELPHEQVFLQPPTSSKPRGPVLPQLELER-----QLGK----- 460
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
343 RSI-----DATD-ELKASK-----SSQKNQESFIGNSQEVEIRITADYQASDEASTSTA 390

QY 461 ----AHLKRAILGOFEALRLHALCVLRE-----VDLLRAVISQTLQRLSKAY-----AELD 508
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
391 NKLALHLKEALASTEERQRLQRCMKRFLKNKASLTENVKVEVHKS-QSELEKFRLOASAKNI 449

QY 509 REDDFCEAAEAADIQKTHQKPEARMPRLSQGKPDIFHRLGPLSVFSAKNRWLVGPVH 568
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
450 EGDGLLERS-----VEASSKFTLSLTGPDFTSHKVKDPFKRLGPIATFSARRHWTAPRCVR 505

QY 569 LTRG-----EGG-----FGLTLRGDSPVLIAAVIFG 594
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
506 LQKGSSLYHSVPSNDKCPDLNDDDEHGGVNYLYKEEFENFGFVRGDPVPIIAHVEIN 565

QY 595 SQAAAGLKEGYIVSVNGQPCRRWRHAEVITELKAAGAGASLQVSVLLPSRLPSLGD 654
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
566 SLADLGKKEGDFIVEIAGVDVKWYSHQVQLIQSCGST-LELRVITPMDRNLKPLSS 624

QY 655 RRPVLLGPRLRLSQRHECKT--PASTWASRPRLNWSRKAQQGKTGCPGCPAPV 709
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
625 K-----GSLSTLSAASSSGISGFFSPSTIAAKPKLHL-----KTSSSSRPAGSV 669
```

RESULT 14

```
Q8N9D6 PRELIMINARY; PRT; 298 AA.
ID Q8N9D6;
AC Q8N9D6;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Hypothetical protein FLJ37682.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hippocampus;
RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K., Isogai T.,
RA "NEDO human cDNA sequencing project.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK095001; BAC04471.1; -.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001478; PDZ.
DR SMART; SM00228; PDZ; 1.
KW Hypothetical protein.
SQ SEQUENCE 298 AA; 33380 MW; B62E0B882DEA00F5 CRC64;

Query Match      10.4%; Score 394; DB 4; Length 298;
Best Local Similarity 40.9%; Pred. No. 3.2e-20;
Matches 83; Conservative 41; Mismatches 77; Indels 2; Gaps 1;

QY 453 ERRQLGKHLKRAILGOFEALRLHALCVLREVDLLRAVISQTLQRLSKAYELDREDD 512
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
15 QORRQLGKSHLRATAHHEESVREASLCKLRISIEVLQKVLCAQAQERSRLTYAQHQEED 74

QY 513 FCEAAEADIQKTHQKPEARMPRLSQGKPDIFHRLGPLSVFSAKNRWLVGPVHLTRG 572
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
75 LNLIDAPSVARTEQEVDIILIPQFSKLTPTDFFQKLGFLSVFSANKRWTTPPSIRFTAE 134

QY 573 EGGFGLTLRGDSPVLIAAVIPGSQAAAGLKEGYIVSVNGQPCRRWRHAEVITELKAAG 632
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
135 EGDGLFTLRAGNAPVQVHFLDPYCSAVAGAREGDYIVSLQVLDCKWLTISEVWKLKLSFG 194

QY 633 EAGASLQVSVLLPSRLPSLGRD 655
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
195 EDEIKMKVSVLLDST--PSMHNK 215
```

RESULT 15

```
Q9P257 PRELIMINARY; PRT; 1421 AA.
ID Q9P257;
AC Q9P257;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein KIAA1471 (Fragment).
GN KIAA1471.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20277482; PubMed=10819331;
RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2004, 12:02:51 ; Search time 30 Seconds
(without alignments)
1244.186 Million cell updates/sec

Title: US-10-697-266-2
Perfect score: 3774
Sequence: 1 MILBERPDGAGBESPR1Q.....QPCAPVKPAPSSLKHPGW 723

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3774	100.0	723	4	US-09-849-334-2
2	3774	100.0	723	4	US-10-274-878-2
3	2331	61.8	634	4	US-09-849-334-4
4	2331	61.8	634	4	US-10-274-878-4
5	268	7.1	921	4	US-09-800-729-199
6	264.5	7.0	868	4	US-09-800-729-106
7	155.5	4.1	942	4	US-08-685-852-3
8	134.5	3.6	450	2	US-08-665-037-2
9	134.5	3.6	450	2	US-08-666-067-2
10	134.5	3.6	450	2	US-08-732-870-2
11	134	3.6	1266	4	US-09-252-991A-30851
12	131.5	3.5	2293	3	US-09-368-590-2
13	129	3.4	751	4	US-09-252-991A-27424
14	128.5	3.4	658	4	US-09-252-991A-18483
15	128	3.4	864	4	US-09-883-096-2
16	125.5	3.3	635	4	US-09-252-991A-18485
17	123.5	3.3	1792	4	US-09-561-818A-4
18	123.5	3.3	1800	4	US-09-561-818A-8
19	123.5	3.3	1816	4	US-09-561-818A-2
20	123.5	3.3	1824	4	US-09-561-818A-6
21	123	3.3	609	4	US-09-252-991A-24893
22	123	3.3	643	4	US-09-252-991A-22490
23	122	3.2	358	4	US-09-740-027-4
24	121.5	3.2	729	4	US-09-252-991A-22350
25	120	3.2	1665	4	US-09-858-664A-2
26	120	3.2	1665	4	US-10-274-978-2
27	120	3.2	6095	3	US-09-144-085-2

28	119.5	3.2	223	4	US-09-124-238A-34	Sequence 34, Appl
29	119.5	3.2	223	4	US-09-721-975-34	Sequence 34, Appl
30	119.5	3.2	223	4	US-09-986-621-34	Sequence 34, Appl
31	119.5	3.2	457	4	US-09-124-238A-10	Sequence 10, Appl
32	119.5	3.2	457	4	US-09-721-975-10	Sequence 10, Appl
33	119.5	3.2	457	4	US-09-986-621-10	Sequence 10, Appl
34	119.5	3.2	819	4	US-09-651-656-15	Sequence 15, Appl
35	119.5	3.2	819	4	US-09-650-855-15	Sequence 15, Appl
36	118	3.1	1445	1	US-08-015-986A-2	Sequence 2, Appl
37	118	3.1	1445	2	US-08-446-363-2	Sequence 2, Appl
38	117.5	3.1	343	4	US-09-252-991A-24630	Sequence 24630, A
39	117.5	3.1	1050	3	US-09-045-632-50	Sequence 50, Appl
40	117.5	3.1	2504	4	US-08-851-567B-12	Sequence 12, Appl
41	117.5	3.1	2504	4	US-09-817-514A-8	Sequence 8, Appl
42	117	3.1	629	4	US-09-252-991A-31243	Sequence 31243, A
43	116.5	3.1	1739	4	US-09-976-594-76	Sequence 76, Appl
44	116	3.1	457	4	US-09-124-238A-1	Sequence 1, Appl
45	116	3.1	457	4	US-09-721-975-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-849-334-2
; Sequence 2, Application US/09849334
; Patent No. 6500655
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001099-CIP
; CURRENT APPLICATION NUMBER: US/09/849,334
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Human
US-09-849-334-2

Query Match	100.0%;	Score	3774;	DB	4;	Length	723;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	723;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	MILBERPDGAGBESPR1Q	SRKPKRTRVSSLRGREGLDVCAWRCRVHRGDEPVR	60			
Db	1	MILBERPDGAGBESPR1Q	SRKPKRTRVSSLRGREGLDVCAWRCRVHRGDEPVR	60			
QY	61	VHVGPMPQLHVGCDLSLTQ	CCGQLOQSRRAIHOQIDKELQMTGAENLYRATSNRVR	120			
Db	61	VHVGPMPQLHVGCDLSLTQ	CCGQLOQSRRAIHOQIDKELQMTGAENLYRATSNRVR	120			
QY	121	ETVALELSYNSNLQLLKBEEL	SGGVDPCRHSSEAVTPVMIPLGLKETELDMSTELK	180			
Db	121	ETVALELSYNSNLQLLKBEEL	SGGVDPCRHSSEAVTPVMIPLGLKETELDMSTELK	180			
QY	181	ELISVHFGEDGASYEAREIREAL	QRMRTSPRNESGLELLTAYYNOICFIDARFLTPAR	240			
Db	181	ELISVHFGEDGASYEAREIREAL	QRMRTSPRNESGLELLTAYYNOICFIDARFLTPAR	240			
QY	241	SLGLFFHWYDSLTVGPAQORALAF	EKGSVLNICALHTQIGARODRSCTEGARRAMEAFQ	300			
Db	241	SLGLFFHWYDSLTVGPAQORALAF	EKGSVLNICALHTQIGARODRSCTEGARRAMEAFQ	300			
QY	301	RAAGAFSLLENFNSHAPSPDMSAAS	LCALEQLMAQAQECVFEGLSPPASMAPQDCLAQL	360			
Db	301	RAAGAFSLLENFNSHAPSPDMSAAS	LCALEQLMAQAQECVFEGLSPPASMAPQDCLAQL	360			
QY	361	PLAQEAQVAAYEELVHRTMAOPV	HHVYVSVTALVHVKAIEYFRSLAHYHVAMALCDGS	420			
Db	361	PLAQEAQVAAYEELVHRTMAOPV	HHVYVSVTALVHVKAIEYFRSLAHYHVAMALCDGS	420			

```
Db      361  RLAEAAQVAAEYRLVHRTWAQPPVHDYVPVSWTALVHVKAAYFRSLAHYHVAMALCDGS 420
Qy      421  PATEGELPTHEQVFLQPTSSKPRGPVLPOELERRQLGKHLKRALIGOEALRLHALC 480
Db      421  PATEGELPTHEQVFLQPTSSKPRGPVLPOELERRQLGKHLKRALIGOEALRLHALC 480
Qy      481  RVLRVDLLRAVISQTLQSLAKYAEUDREDDFCEAAEAPDIQPKTHQKPEARMPLRSQ 540
Db      481  RVLRVDLLRAVISQTLQSLAKYAEUDREDDFCEAAEAPDIQPKTHQKPEARMPLRSQ 540
Qy      541  KGPDIHRLGFLSVFSAKNRWLVGPHVLTGRGEGFGLTLRGDSPVLIAAVIPGSQAAA 600
Db      541  KGPDIHRLGFLSVFSAKNRWLVGPHVLTGRGEGFGLTLRGDSPVLIAAVIPGSQAAA 600
Qy      601  GLKEGDYIVSVNGQPCRWHRHAEVVTELKAAGEAGASIQVVSLLPSSRLPSLGDRRPVLL 660
Db      601  GLKEGDYIVSVNGQPCRWHRHAEVVTELKAAGEAGASIQVVSLLPSSRLPSLGDRRPVLL 660
Qy      661  GPRGLLSQREHGCKTTPASTWASPRLLNWSRKAQOGKTCGCPQPCAPVKPAPSSLKHP 720
Db      661  GPRGLLSQREHGCKTTPASTWASPRLLNWSRKAQOGKTCGCPQPCAPVKPAPSSLKHP 720
Qy      721  GWP 723
Db      721  GWP 723

RESULT 2
US-10-274-878-2
; Sequence 2, Application US/10274878
; Patent No. 6670163
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001099-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,878
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Human
US-10-274-878-2

Query Match 100.0%; Score 3774; DB 4; Length 723;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MILBERPDGAGAESPRLOISRRKPRKTRVSSLRGREGRLDVCAMRGCRVHRGEDPVR 60
Db      1  MILBERPDGAGAESPRLOISRRKPRKTRVSSLRGREGRLDVCAMRGCRVHRGEDPVR 60
Qy      61  VHVGMNPMQLHAVGCDSTQIQCGLOQSRRAQIHOQIDKELQMTGAENLYRATSNRVR 120
Db      61  VHVGMNPMQLHAVGCDSTQIQCGLOQSRRAQIHOQIDKELQMTGAENLYRATSNRVR 120
Qy      121  ETVALELSYVNSNLQLLKEELEELSGGVDPGHSGEAVTVPMIPLGLKETKELDWSPLK 180
Db      121  ETVALELSYVNSNLQLLKEELEELSGGVDPGHSGEAVTVPMIPLGLKETKELDWSPLK 180
Qy      181  ELISVHFGEKGASYPAEIRELEALQAMRTPSRNSSLGELLTAYYNQLCFDLARELTAR 240
Db      181  ELISVHFGEKGASYPAEIRELEALQAMRTPSRNSSLGELLTAYYNQLCFDLARELTAR 240
Qy      241  SLGLFFHYVDSLTVGVAQORALAFKGSVLFNIGALHTQIGARQDRSCTSGARRAMEAFQ 300
```

```
Db      241  SLGLFFHYVDSLTVGVAQORALAFKGSVLFNIGALHTQIGARQDRSCTSGARRAMEAFQ 300
Qy      301  RAAGAFSLLENFSAFSPDMASAASICALQOLMAQAQECVFEGLSPPASWAPODCLAQ 360
Db      301  RAAGAFSLLENFSAFSPDMASAASICALQOLMAQAQECVFEGLSPPASWAPODCLAQ 360
Qy      361  RLAEAAQVAAEYRLVHRTWAQPPVHDYVPVSWTALVHVKAAYFRSLAHYHVAMALCDGS 420
Db      361  RLAEAAQVAAEYRLVHRTWAQPPVHDYVPVSWTALVHVKAAYFRSLAHYHVAMALCDGS 420
Qy      421  PATEGELPTHEQVFLQPTSSKPRGPVLPOELERRQLGKHLKRALIGOEALRLHALC 480
Db      421  PATEGELPTHEQVFLQPTSSKPRGPVLPOELERRQLGKHLKRALIGOEALRLHALC 480
Qy      481  RVLRVDLLRAVISQTLQSLAKYAEUDREDDFCEAAEAPDIQPKTHQKPEARMPLRSQ 540
Db      481  RVLRVDLLRAVISQTLQSLAKYAEUDREDDFCEAAEAPDIQPKTHQKPEARMPLRSQ 540
Qy      541  KGPDIHRLGFLSVFSAKNRWLVGPHVLTGRGEGFGLTLRGDSPVLIAAVIPGSQAAA 600
Db      541  KGPDIHRLGFLSVFSAKNRWLVGPHVLTGRGEGFGLTLRGDSPVLIAAVIPGSQAAA 600
Qy      601  GLKEGDYIVSVNGQPCRWHRHAEVVTELKAAGEAGASIQVVSLLPSSRLPSLGDRRPVLL 660
Db      601  GLKEGDYIVSVNGQPCRWHRHAEVVTELKAAGEAGASIQVVSLLPSSRLPSLGDRRPVLL 660
Qy      661  GPRGLLSQREHGCKTTPASTWASPRLLNWSRKAQOGKTCGCPQPCAPVKPAPSSLKHP 720
Db      661  GPRGLLSQREHGCKTTPASTWASPRLLNWSRKAQOGKTCGCPQPCAPVKPAPSSLKHP 720
Qy      721  GWP 723
Db      721  GWP 723

RESULT 3
US-09-849-334-4
; Sequence 4, Application US/09849334
; Patent No. 6500655
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001099-CIP
; CURRENT APPLICATION NUMBER: US/09/849,334
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-849-334-4

Query Match 61.8%; Score 2331; DB 4; Length 634;
Best Local Similarity 67.0%; Pred. No. 3.3e-210;
Matches 473; Conservative 49; Mismatches 112; Indels 72; Gaps 6;

Qy      1  MILBERPDGAGAESPRLOISRRKPRKTRVSSLRGREGRLDVCAMRGCRVHRGEDPVR 60
Db      1  MILBERPDGAGAESPRLOISRRKPRKTRVSSLRGREGRLDVCAMRGCRVHRGEDPVR 60
Qy      61  VHVGMNPMQLHAVGCDSTQIQCGLOQSRRAQIHOQIDKELQMTGAENLYRATSNRVR 120
Db      27  -----KGYGSFVQNPQGLQSHRARLHQOISKELMRMTGAENLYRATSNRVR 74
Qy      121  ETVALELSYVNSNLQLLKEELEELSGGVDPGHSGEAVTVPMIPLGLKETKELDWSPLK 180
Db      75  ETVALELSYVNSNLQLLKEELEELSGVDVQPEGEGITIPMIPGLKETKELDWSPLK 134
Qy      181  ELISVHFGEKGASYPAEIRELEALQAMRTPSRNSSLGELLTAYYNQLCFDLARELTAR 240
```

Db 135 ELISHFGEDGTSFETEIQELEDLRQATRTPSRDEAGLDLLAAAYSQLCFDARFFSPSR 194
QY 241 SLGLFFHYDLSLTGVPAQORALAFKSGSVLFGALHTQIGARQDRSCTEGARRAMEAFQ 300
Db 195 SPGLLFHYDLSLTGVPAQORALAFKSGSVLFGALHTQIGARQDRSCTEGTHAAEAFAQ 254
QY 301 RAAGAFSLLENFNSHAPSDFMSAASLSLEQMLIAQAQECIFKGLLLPASATPDCPDOL 360
Db 255 RAAGAFSLLENFNSHAPSDFMSAASLSLEQMLIAQAQECIFKGLLLPASATPDCPDOL 314
QY 361 RLAEAAQAAEAYEYRLVHRTMAQPPVHDYVPSWTALVHVKAHYFRSLAHYHVMALCDGS 420
Db 315 QLAEEAAQVATEYGLVHRMAQPPVRDYLPSWTNLAVHKAHFCALAHYHVMALCESH 374
QY 421 PATEGELPTHQVFLQPTSSKPRGVPVLPQELERRQLGKAHLKRAILGQBEALRLHALC 480
Db 375 PA-KGELARQEHVF-QPSTPHEPLGPTLPQHPEDRRKLAKAHLKRAILGQBEALRLHALC 432
QY 481 RVLREVDLLRAVISOTLORSIAKYAELDRDDFCEAAEAPDIQPKTHOKPEARMPRLSQG 540
Db 433 RVLKVDLLQVVVTVQALRRSLAKYSLQEREDDFEATEAPDIQPKTHOTPE----- 483
QY 541 KGPDIHRLGLSPVFSAKNRWLVGPVHLTRGEGFGTLRGDSFVLIAAIVPGSQAAAA 600
Db 484 -----GPLSVFSTKNRWQLVGPVHMTREGGFGTLRGDSFVLIAAIVPGSQAES 534
QY 601 GLKGDYIVSVNGQPCRWVHAEVVTTELKAAAGEAGASLOVLSLPSRLPSIGDRRPVLL 660
Db 535 GLKGDYIVSVNGQPCRWVHAEVVTTELKAAAGEAGASLOVLSLPSRLPSIGDRRPVLL 594
QY 661 GPRGLLRQREHGCKTPASTWASPRPLNWSKKAQGGTGGCQPC 706
Db 595 W-----NQREGFETPMTTRTPWPLILGWSKKNKQGGTGGHPDPC 634

RESULT 4

US-10-274-878-4
; Sequence 4, Application US/10274878
; Patent No. 6670163
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO01099-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,878
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-274-878-4

Query Match 61.8%; Score 2331; DB 4; Length 634;
Best Local Similarity 67.0%; Pred. No. 3.3e-210;
Matches 473; Conservative 49; Mismatches 112; Indels 72; Gaps 6;
QY 1 MILERPDGAGAGESPRLQISRRKPRKTRVSSLRGREGRLDVCAMGRVHRGEDPVR 60
Db 1 MILERPDGQGTGESSRPQDD-----GSIR----- 26
QY 61 VHVGPWNLHAGVCDLSLTQCGLOSRRAHQIDKELQMRGCAENLYRATSNRVR 120
Db 27 -----KGVGSFVQNPQGLQSHRLHQOISKELSMRGTCAENLYRATSNRVR 74
QY 121 ETVALELSVNSNLQLLKEELAEELSGGVDPPGRHGSEAVTVPMPLGLKXETKELDWSTPLK 180

Db 75 ETVALELSVNSNLQLLKEELAEELSGGVDPPGRHGSEAVTVPMPLGLKXETKELDWSTPLK 134
QY 181 ELISVHFGEAGSYBAEIRELEALRQAMRTPSRNBESGLELLTAYYNQLCFDARFLTPAR 240
Db 135 ELISHFGEDGTSFETEIQELEDLRQATRTPSRDEAGLDLLAAAYSQLCFDARFFSPSR 194
QY 241 SLGLFFHYDLSLTGVPAQORALAFKSGSVLFGALHTQIGARQDRSCTEGARRAMEAFQ 300
Db 195 SPGLLFHYDLSLTGVPAQORALAFKSGSVLFGALHTQIGARQDRSCTEGTHAAEAFAQ 254
QY 301 RAAGAFSLLENFNSHAPSDFMSAASLSLEQMLIAQAQECIFKGLLLPASATPDCPDOL 360
Db 255 RAAGAFSLLENFNSHAPSDFMSAASLSLEQMLIAQAQECIFKGLLLPASATPDCPDOL 314
QY 361 RLAEAAQAAEAYEYRLVHRTMAQPPVHDYVPSWTALVHVKAHYFRSLAHYHVMALCDGS 420
Db 315 QLAEEAAQVATEYGLVHRMAQPPVRDYLPSWTNLAVHKAHFCALAHYHVMALCESH 374
QY 421 PATEGELPTHQVFLQPTSSKPRGVPVLPQELERRQLGKAHLKRAILGQBEALRLHALC 480
Db 375 PA-KGELARQEHVF-QPSTPHEPLGPTLPQHPEDRRKLAKAHLKRAILGQBEALRLHALC 432
QY 481 RVLREVDLLRAVISOTLORSIAKYAELDRDDFCEAAEAPDIQPKTHOKPEARMPRLSQG 540
Db 433 RVLKVDLLQVVVTVQALRRSLAKYSLQEREDDFEATEAPDIQPKTHOTPE----- 483
QY 541 KGPDIHRLGLSPVFSAKNRWLVGPVHLTRGEGFGTLRGDSFVLIAAIVPGSQAAAA 600
Db 484 -----GPLSVFSTKNRWQLVGPVHMTREGGFGTLRGDSFVLIAAIVPGSQAES 534
QY 601 GLKGDYIVSVNGQPCRWVHAEVVTTELKAAAGEAGASLOVLSLPSRLPSIGDRRPVLL 660
Db 535 GLKGDYIVSVNGQPCRWVHAEVVTTELKAAAGEAGASLOVLSLPSRLPSIGDRRPVLL 594
QY 661 GPRGLLRQREHGCKTPASTWASPRPLNWSKKAQGGTGGCQPC 706
Db 595 W-----NQREGFETPMTTRTPWPLILGWSKKNKQGGTGGHPDPC 634

RESULT 5

US-09-800-729-199
; Sequence 199, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 199
; LENGTH: 921
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (247)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (362)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (603)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-800-729-199

Query Match 7.1%; Score 268; DB 4; Length 921;
Best Local Similarity 26.7%; Pred. No. 6.4e-16;

Db 114 ----GAGGTCSATNLSRVAGLEKQLAIELKVKQGAENMIQTYSNGSTKDRKLLLTAAQM 169
QY 243 GUFFHWYDSLT-----GVPAQRALAFEGKSVLFTNGALHTQIGA-----282
Db 170 ----LQDSKTKIDIIIRMQLRRALQAGLENQAAPDDTQOSP--DLGAVELRIEELRHFF 222
QY 283 RODRSCTEGARAMEAFQAAAGAFSLLENFENSHAPSPDMSAASLCALEQLMMAQAQECVF 342
Db 223 RVEHAVAEGAKVRL-----LSNAKAPDRKAVSEAQEKUTESQKLGILLR 268
QY 343 EGLSPSPASMAPQCLQRLAQAQAAVAAYRULVHRTMAQPPVHDY-----VPVSWTALVH 398
Db 269 EALERRLGLPAD-HPKGLRLLEELAASSAASFSTRLAGFPFATHYSTLCKPAPLTGTL 327
QY 399 VKAEYERSLAH-----YHVAMALCDGSPATEGELPTHEQVFLQPP-----438
Db 328 VRVVGCRDLPETIPMNPFTSM--GGFGT----PDSRPFLSRPARGLYSRSGSLSGRSSU 381
QY 439 ----TSSKPRGP-----VLPQELERERQLGKAHLKRAILGQ 470
Db 382 KAAEANTSEVSTVLKLDNTVVGTSWKPCGNWDQSFTELEERARE-----LELAVFWR 436
QY 471 BEALRLHALC--RVLEVDLLRAVISQTLQRSLSAKYAEILDREDDFCEAAEAPDIOPKTHQ 528
Db 437 DQ----RGLCALFKLEDFL-----DNERHEVQLDMREPQGCCLVAEV-----TFR 477
QY 529 KPE-ARMPRL-----SQKGPDIFHRLGPLSVFSAKNRWLVGPFVHLTR-----571
Db 478 NVIEIRIPRLRQKKIFSKQOGKA---FORAQMNIDVA--TW-----VRLLERLIPNAT 527
QY 572 GEGGGLTLRGDSPVLIAAVIPGSOAAAAGLKEGDYIVSVNGQPCRWHRHAEVVTLEKAA 631
Db 528 GNGTF-----SP-----CASPGSEARTG-----DISVE-----KING 556
QY 632 GEAGASLQVSVLLPSRLSDRRPVLGPRGLLRSQREHGCKTPASTWASPRPLNWS 691
Db 557 TDSQSPQKSRDPSPSSLS--PI-----QBSTAPELPSETQETPGPAL----601
QY 692 RKAQOGKGGCPQCAPVKPAP 713
Db 602 -----CSPLRKSP 609

RESULT 8

US-08-665-037-2
; Sequence 2, Application US/08665037
; Patent No. 5895813

; GENERAL INFORMATION:
; APPLICANT: Seedorf, Klaus
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF TKA-1 RELATED
; TITLE OF INVENTION: DISORDERS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,037
; FILING DATE: June 13, 1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/005,167
; FILING DATE: October 13, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 220/156
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 450 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-665-037-2

Query Match 3.6%; Score 134.5; DB 2; Length 450;
Best Local Similarity 24.6%; Pred. NO. 0.00071;
Matches 84; Conservative 21; Mismatches 111; Indels 125; Gaps 14;

QY 486 VLLRAVISQTLQRSLSAKYAEILD-----REDDFCEAAEAPDIOPKTHQKPEARNPRLSQCK 541
Db 72 VQRIKAVEGT--RLVVDQETDEELRRQLCTTEMAQORGLPPADHPWEPK-----121
QY 542 GPDIFHRLGPLSVFSAKNRWLVGPV-----HITRGGEGGLTIRGDSF---VLIAA 590
Db 122 -FDWAHTGSHSSEAGKD--VSGPLRELRLPRLCHLRKGPQGYGNLHSDKSRPGQYIRS 177
QY 591 VIPGQAAAAGLKEGDYIVSVNGQPCRWHRHAEVVTLEKAAAGEAGASLQVV-----641
Db 178 VDPGFGARGSLRAQDRLIEVNGQVGLRHAENVASIKAR-EDEARLLVDPETDHEFK 236
QY 642 -----SLGPS-----SRLP-----SLGDRP---657
Db 237 RLVRVPTSEHVGEPLSPVYNTGTSAPQLNGGSACSRSDLPQSDKDTEDGSAWKQDPQOE 296
QY 658 --VLLGPRGLLRSQREHGCKT-----PASTWAS 683
Db 297 SGLHLSPTAAEAARRRLEPCESTSARHRWTGTSVKSSATSEPLPACLGTLGLPLHGPWAS 356
QY 684 RPPLNWSRKAQOGKGGCPQCAPVKPA---PPSSLKHGP 721
Db 357 ACPEL-----PQOWTGMSCHCPEISPSGEPSPCPCPPG 392

RESULT 9

US-08-666-067-2
; Sequence 2, Application US/08666067
; Patent No. 5922842
; GENERAL INFORMATION:
; APPLICANT: Seedorf, Klaus
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF TKA-1 RELATED
; TITLE OF INVENTION: DISORDERS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/666,067
FILING DATE: June 13, 1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/005,421
FILING DATE: October 13, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 220/157
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-666-067-2

Query Match 3.6%; Score 134.5; DB 2; Length 450;
Best Local Similarity 24.6%; Pred. No. 0.00071;
Matches 84; Conservative 21; Mismatches 111; Indels 125; Gaps 14;
QY 486 VDLRAVISQTLQSLAKYAEID---REDDFCEAAEAPDIQPKTHQKPEARMPLRSQK 541
Db 72 VQRIKAVEGQT--RLVVVDQETDEELRRQLTCTEEMAQRLGLPPAHDPWEK-----121
QY 542 GPDIFHRLGLPLSVFSAKNRMVLGVF-----HLTRGEGGGLTLRGDSP---VLIAA 590
Db 122 -PDWAHTGSHSSEAGKXD---VSGPLRELRLPRLCHLRKPGQYGFNLHSDKSRPQYIRS 177
QY 591 VIPGSQAAAGLKEGDYIVSVNGQPCRWRAEAVVTELKAAGEAGASIQV-----641
Db 178 VDPGSPGARSGLRAQDRLLIEVNGQNVGLRHAENVASIKAR-EDEARLLVVDPTDEHFK 236
QY 642 -----SLLPS-----SRLP-----SLGDRRP---657
Db 237 RLVRTPTTEHVEGGLPLSPVNTGTPAQLNGGSACSSRDLFGSKDXTEDGSAMKQDPQ 296
QY 658 --VLLGPRGLLRSOREHGCKT-----PASTWAS 683
Db 297 SGLHLSPTAAEARRRLEPCESTSARHRTGTGSKSSATSBPLPACLTGLGLPHGPWAS 356
QY 684 PRPLNWSRKAQCKTGCPQPCAPVKA---PPSSLKHG 721
Db 357 ACPEL-----PQOWTGGWSCHCEISPSGPEPPSCPCPPG 392

RESULT 10
US-08-732-870-2
Sequence 2, Application US/08732870
Patent No. 5945523
GENERAL INFORMATION:
APPLICANT: Ullrich, Axel
APPLICANT: Seedorf, Luitgard
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT
TITLE OF INVENTION: OF TKA-1 RELATED
TITLE OF INVENTION: DISORDERS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: Lyon & Lyon
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage

COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/732,870
FILING DATE: October 15, 1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/666,037
FILING DATE: June 13, 1996
APPLICATION NUMBER: 08/666,067
FILING DATE: June 13, 1996
APPLICATION NUMBER: 60/005,167
FILING DATE: October 13, 1995
APPLICATION NUMBER: 60/005,421
FILING DATE: October 13, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 222/247
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-732-870-2
Query Match 3.6%; Score 134.5; DB 2; Length 450;
Best Local Similarity 24.6%; Pred. No. 0.00071;
Matches 84; Conservative 21; Mismatches 111; Indels 125; Gaps 14;
QY 486 VDLRAVISQTLQSLAKYAEID---REDDFCEAAEAPDIQPKTHQKPEARMPLRSQK 541
Db 72 VQRIKAVEGQT--RLVVVDQETDEELRRQLTCTEEMAQRLGLPPAHDPWEK-----121
QY 542 GPDIFHRLGLPLSVFSAKNRMVLGVF-----HLTRGEGGGLTLRGDSP---VLIAA 590
Db 122 -PDWAHTGSHSSEAGKXD---VSGPLRELRLPRLCHLRKPGQYGFNLHSDKSRPQYIRS 177
QY 591 VIPGSQAAAGLKEGDYIVSVNGQPCRWRAEAVVTELKAAGEAGASIQV-----641
Db 178 VDPGSPGARSGLRAQDRLLIEVNGQNVGLRHAENVASIKAR-EDEARLLVVDPTDEHFK 236
QY 642 -----SLLPS-----SRLP-----SLGDRRP---657
Db 237 RLVRTPTTEHVEGGLPLSPVNTGTPAQLNGGSACSSRDLFGSKDXTEDGSAMKQDPQ 296
QY 658 --VLLGPRGLLRSOREHGCKT-----PASTWAS 683
Db 297 SGLHLSPTAAEARRRLEPCESTSARHRTGTGSKSSATSBPLPACLTGLGLPHGPWAS 356
QY 684 PRPLNWSRKAQCKTGCPQPCAPVKA---PPSSLKHG 721
Db 357 ACPEL-----PQOWTGGWSCHCEISPSGPEPPSCPCPPG 392
RESULT 11
US-09-252-991A-30851
Sequence 30851, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788

```
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30851
; LENGTH: 1266
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30851
```

```
Query Match          3.6%; Score 134; DB 4; Length 1266;
Best Local Similarity 22.1%; Pred. No. 0.0043;
Matches 178; Conservative 81; Mismatches 288; Indels 260; Gaps 44;
```

```
QY 3 LRRPDGAGAESPLQISRRKPRKTRVSSLR-GRREGIRDVCAWRCGRVIRGEDPVV 61
Db 421 IQRFLGEPGNPHGLQPGAR--RRVRAGMKRQRREHLAD-----HRGED--GR 467

QY 62 HVGPMNPOLHAVGCDLSLTQICGQQLGRRRAHQHIDKELQMTGAENLYR----- 112
Db 468 HGGGRQ-----GC--LRPRAGLDLRR--VHQRREGGQGGAGRRRRHHGDAGAGL 516

QY 113 ----ATSNRVRTVALELSYNSNLIQLKELEELSGGVDPRHSGSEAVTVMPIPLGLKE 169
Db 517 FLQAAARDRAFPQGRQHRPAGDGLQ----QPADLPERRHPGH-----FGFPGLRLE 564

QY 170 TKELDWSTPLKELISVHFCEDGASYEATRELEALRQAMRT-PSRNESGLELLTAYYNOL 228
Db 565 HRLLOGQ--LRHPPLHRRTPQGG-----RALHPLRARRRRFGKRGGG----- 606

QY 229 CFLDARFATPARSLGFFHWDYSLTGVPAAQR-----ALAFKGSVLFNT----- 273
Db 607 -----CFGLDLG-----NVQVPEGRDRDHLPPGGRALRGDADLRVADADPPSR 651

QY 274 ----GALHTQIGA-----RODRCTEGARPAEAFQAAGAFSILLRNFESHAPSDFS 322
Db 652 RPSRFGVHQALRATERSRQRPHSTAPGARRRRGARRA-----DHGQGPQP 700

QY 323 AASLCALQMLMAQAQECVFEGLSPPAS-----MAPQDCLAQLRLAQEAQVAAYRVLVH 377
Db 701 AG--AAGHRLVADAPP-----SPRSPGSGRRGP-----AAEHS-VH 734

QY 378 RTMAQPPV----HDYVPVSWTA-----LVHVKAQVFRSLAHYHVAMALCDGSPATEGELP 428
Db 735 RPQGGPMPSILGHNVGGARSAGNLLLSLDSDSSEALPYAFV-----QATEAEV- 785

QY 429 THEQVLPQPTSSKPRGPVLP--QELEERRQLGKAHLKRAILGQBEALRLHALCRVLRV 486
Db 786 -----DAAAARAAERAYPHYRLSATRRAG--FLEATASRLDALGDDFVALVRRET 833

QY 487 DLLRAVISQTLQRS---LAKYAELEDDDDCEAAEAPDIQ-KTHQKPEARNPRLSQGKG 542
Db 834 ALPAARIOGERTITANQRLFAEVLRRGDF-HGAIDRGQPGRTPPRPLRQWRIG---- 888

QY 543 PDIFHRLGPLSVFSAKNRWLVGPVHLTRGEGFGLTLRGDSPVLIAAVIPGSQAAAAGL 602
Db 889 ----LGPVAVFGASN-----FPLAFSTAGDSAAALAAAGCPVVXA----- 925

QY 603 KEGDYIVSVNGPCRWWRHAEVWTE--LKAGEAGASLOVSLSPSSRLPSLGRRPVL- 659
Db 926 -HGHHMAT-----AECVADAILQAAADSGMPAGVFNMVYGVSGVEALVHHPAIR 973

QY 660 -LGPRLRLRSOREHCKTPTASTWASPRLLNWS-----RKAQO----- 696
Db 974 AVGFTGSLKGRAL--CDLAA--ARPQIPVFAEMSSINPLVLPEALRRRGRQVAEELA 1029

QY 697 -GKTGGCPQCAPVKPAPPSSLKHPGW 722
Db 1030 ASVTILGCGQFC--TKPGLVLGRSPGF 1054
```

RESULT 12

US-09-368-590-2

```
; Sequence 2, Application US/09368590
; Patent No. 6187563
; GENERAL INFORMATION:
; APPLICANT: Solimena, Michele
; TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
; TITLE OF INVENTION: AUTOANTIGENS OF AUTOIMMUNE DISEASES
; FILE REFERENCE: 101918-200 (OCR-941)
; CURRENT APPLICATION NUMBER: US/09/368,590
; EARLIER FILING DATE: 1998-08-04
; EARLIER APPLICATION NUMBER: 60/095,657
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2293
; TYPE: PRT
; ORGANISM: Human
US-09-368-590-2
```

```
Query Match          3.5%; Score 131.5; DB 3; Length 2293;
Best Local Similarity 21.2%; Pred. No. 0.019;
Matches 161; Conservative 102; Mismatches 296; Indels 201; Gaps 38;
```

QY	29	TRVSSLRGERREGIRDVCAWRCGRVIRGEDPVVHVGPMNPOLHAVGCDLSLTQICGQLOS	88
Db	1529	TAAATMAEKQGLNE--AWAELELMG---TRAQLLAASRELHKFFSDA-RELQ-GQIEE	1581
QY	89	RAAQHQHIDKELQMTGAENLYRATSNRVRTVALELSYNSNLIQLKELEELSGG-	147
Db	1582	KERRLP-----RLTTPPEPSPASSMQRT-----LRAFEHDLQLLVSVQRLQEGA	1627
QY	148	-----VDPGRHSGSEAVTVMPIPLGKETKELDWSTPLKELIS-----VHFGEDGASY	194
Db	1628	AOLRTVYVAGEH-AEAIA-----SREQEVLQGWKELLSACEDARLHVSTADALRF	1676
QY	195	EAEIRELEALRQAMRT-----PSRNESGLELLTAYYNOL-CFLDARF--LTPARSLGLF	245
Db	1677	HSQVRDLLSWMGDGIASQIGAADKPRDVSSEVLMNYHQGLKTELEARVPELTTCQELG--	1734
QY	246	FHWYDSLGTGVPAAQRALAFKGSVLFNIGALHTQIGARQDRSCTEGARRAEAFQRAAGA	305
Db	1735	-----RSLINKSAMADEITQALDKLGTKEEVSERKWDHRH-WEWLQOMLEHVHQAQEA	1786
QY	306	PSLIRENFESHAPSDFMSA-----ASLCALQMLMAQAQECVFEGLSPPASMAPODCLAQLR	361
Db	1787	--VVADAMLTQAEPLQLQSRELGSVDVEQLIRRH--AFR---KAAAWEERFSSLR	1837
QY	362	LAQEAQAQVAAYEYVRLVHRTMAQPPVHDYVPSWTALVHVKAQVFRSLAHYHVAMALCDGSP	421
Db	1838	RLTTTEKIKAE-----QSKOPP-----TPLLGRK--FF-----GDP	1866
QY	422	ATEGELPTHEQVFLQPPTSSKPRGPV-----LPOELEERRQLGKAHL	463
Db	1867	T---ELAAKAAPLLRPGYVERGLEFLARRASDTLSAEVTRTVGVYRQELKEERLQPRIDR	1923
QY	464	KRAILGQEEALRLHALCRVLRREVLLRAVISQT-----LQSLAKYAELEDDDDFCEAAE	518
Db	1924	LPEIPGRVEPAALPAAPEDAETATPAABQVPRPERQESADRAEELPRR-----R	1976
QY	519	APDIQPTKHXKPEA--RMPPLSQGKGPDIPIHRLGPLSVFSAKNRWLVGPVHLTRGEGG	575
Db	1977	RPEQESVDQSEEAARRRRRPERQESAHEAAHSL-TLGRYEQMERRRRRRERLERQESS	2035
QY	576	-FGLTLRGDSPVLIAAVIPGSQAAAAAGLKEGDYIVSVNGPCRWWRHAEVWTE--KAA	632
Db	2036	EQEMPIRGD-----LVK GKATL-----ADIVEQLQOEKAG	2065
QY	633	---FAGASLOVSLSPSSRLPS---LGDRRPVLGPRGLLRFSQREHCKTPTASTWASPRP	686
Db	2066	PGLPAGFSLPQPRELPPGRLPNGLPELPTPRPDRPRARORPKPRR-----RPRP	2115
QY	687	LL-----NWSKAAQOQKTCGCPQPCAPVKPAPPSSLKHPGW	722

Db 2116 REGGEGGSRSSAPAQGSAPP--PPPTHVQHEGF 2153

RESULT 13

US-09-252-991A-27424

; Sequence 27424, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 27424

; LENGTH: 751

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-27424

Query Match 3.4%; Score 129; DB 4; Length 751;

Best Local Similarity 21.9%; Pred. No. 0.0054;

Matches 185; Conservative 80; Mismatches 254; Indels 326; Gaps 47;

Qy 28 KTRVSSLRGR---REGLRDVCAMRGCRVHRGDPVRVH-----VGMNPNQLHAVGCD 76

Db 33 KARLDSQERSQASSLDVLG-----DAIRLERELADMKVGDRGVAEHA----- 78

Qy 77 SLTQCCQLOSRAQHQIDKELQMTGAENLYRATSNRVRRTVALELSYVNSNQL 136

Db 79 -----RQLGEROQL-----KRLGIE--ARAAGDAYAR----- 104

Qy 137 LKEELERLSCGVDPGRHSGEAVTPMIPGLKTKETKELDWSTPLKELIS---VHFGEDCAS 193

Db 105 -----LGMQRLD-----MQVRGLRQEQASQAMPLASAFGLVVEASKTAAG 148

Qy 194 YEATIRELEALRQAM---RTP-----SRNESGLELLTAYYNOLCFIDAR---FLPPA 239

Db 149 YQARLRLD--ALRNGLDVGRPALASLIQDSANOSGLG--RTATLDMLEHLNATGMGFAAAQ 206

Qy 240 RSLGLF--FHWYDSLITG--VPAQORALAFKGSVLFNIGALHTQIGARQDRSCTEG-ARR 294

Db 207 MNMGLAGRFQGGIASAFAVAGLVRLALQAQGS-----DSPEQLSASLDRLVLKGRV 260

Qy 295 AMEAFQRAAGAFSLLENFSPHAPSMDMSAASLCALQMLMAQAQBCV----- 341

Db 261 GSEALARLRLPALLSALGNAGEATAGDVG-----ALGALLELOAKNTTPDKADVRMKAWLE 315

Qy 342 FEGLSPPASMAPQDC---LAQIR-----LAQBAQVAAYEYR---LVHRTMAQPPVHDYV 389

Db 316 FVGSGLKRAYQDYDRDLFALRKDGASLLEALNLELAARYDKGKLSAGVASPALEAY- 374

Qy 390 PYSWTALVHVKAEYFRSLAHYHVAMALCDGSPATEGELPTHQVLEPPTSS-----K 442

Db 375 -----RASGEF-----QGLLESQSSVSGSERDA 399

Qy 443 PRGPVLPQEL-----EBRRQ--LGKA-----HLKR--AILGOBEALRLHALCRVLR- 484

Db 400 QRRKMSQELWKASSDSWEAQTALGSALNPYLDNLAKGSVILGESTAELEAYPRTTAG 459

Qy 485 -----EVDLLR-----AVISQTLQSLAKY---AELDR--- 509

Db 460 LTAAGAVLSGLYAKGGRGAIDVLRGRLGRGTAAVGLDTERGAGRVSGGSEIQRVFW 519

Qy 510 -----EDFCEAAEPDIQPKTHQKPEARMPRLSQGKGPDIFFHRLGLPLSVFSKRNWR 562

Db 520 TNWVPVGGDSLTESARRP-----AQRKQTPRRKRGKGGGLKARSLPSLGSFA----- 568

Qy 563 LVGPVHLTRGEGGFG-----LTLRGDSPVLIA-AVIFGSQAAAAAGLKEGDYIVS 610

Db 569 -----GGGLGAMAGKLPRLSRLPTR-NAPLQVASSLIDVAEYVSDLSSEKTI- 616

Qy 611 VNGQPCRWWEHAEVVTTELKAAGAGASLQVSSLLPSSRLPSLGDRRPV---LLGPRGLLR 667

Db 617 -----AYGEAGASL-AGSLAGAGALGASIGSVVPPVVGTLIG--GLVG 654

Qy 668 SQREHGCKTPASTWASPRPLNWSRKAQCKTG---GCP-----QPCAPVKEPAPSSSL 717

Db 655 G-----AIGANGGSELGRLGRLSLAGDPPAASDNKPAVPVPOAEFVAA 697

Qy 718 KHPCW 722

Db 698 V-PNW 701

RESULT 14

US-09-252-991A-18483

; Sequence 18483, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 18483

; LENGTH: 658

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-18483

Query Match 3.4%; Score 128.5; DB 4; Length 658;

Best Local Similarity 20.9%; Pred. No. 0.0048;

Matches 158; Conservative 84; Mismatches 300; Indels 215; Gaps 33;

Qy 7 PDGAGAGES-----PRLQ---ISRRK---PRKTRV-----SSLRGRREGRLDVCANRG 49

Db 15 PPAATAGDRSSVEDPHRPHPLQAALSQRNVLPEQHRAILFPFVGLQGRQ-LAVLVALAG 73

Qy 50 CRVHRGEDPVRVHVGMNPNQLHAVGCDSLTQCCQLOSRAQHQIDKELQ--MRTG- 106

Db 74 RRIQGEDAAGTGLG--HDELDRADADAPPV---LGERGAAFHQVGTGAQHLQRPAD 127

Qy 107 --AENLYRATSNRVRRTVA---LELSYVNSNLQLLKEELEELSGVDPGRHGSSEAVTV 161

Db 128 PLAEEVVERSLADQOQRETVGEMHLLPRFVAGIDLPRLRL---GNVHQARRAAEAEARQP 183

Qy 162 MIPGLKTKETKELDWSTPLKELISVHFG---EDGASYAEIPELEALQAMR--TPSRNE 215

Db 184 -VRGVAGIRRLQGIAPDRDATQVGVGCRYFEAQHGHPRIQFQDLRRRPARLDTFQURQ 242

Qy 216 SGLELLTAYYNQLCFDLARFLTPARSLGLFHHWYDSLTVGPVPAQORALAFKGSVLFNIGA 275

Db 243 OALGETS-----GFGVHRHPASVADHAQR-----LERDA 272

Qy 276 LHTQIGARQDRSCTEGARRAMEAFORAAGAFSLLENFSPHAPSMDMSAASLCALQMLMA 335

Db 273 LQGLGGRLDRITAVQ--RGETEHGTSILGGKMLL---YGVAGKADKTOTPLQAADQ--- 323

Qy 336 QAQECVFEGLSPASNAQDCLAQLRAQEARAAVAEYELVHRTMAQPPVHDYVVPVSWTA 395

Db 324 -----SAPPSV-----LARRLAQQRRTGRQVHAQARAGNRRTQLASFQRL 363

Qy 396 LVHVKAEYFRSL-----AHYHVAMALCDGSPATEGELPTHQVFLQ----- 436

```
Db 364 AVEIDVETFRRLHFQWHPRADEAQSHTVATB-QGAQGTDDHVLRRRQDAHVQAAVHVS 422
QY 437 -----PPTSSKRGVPLPOELE-ERRQLGKAHL-KRAILGQBEALRLHALCRV 482
Db 423 LRTQYVPAAGLAPVTDQAGEELAFPLEVGAETALGMEFGKTGQPGQVGGTAENLLQV 482
QY 483 LREV--DLLRAVISQTLQSLAKY-AELDREDDFCEAAE-----APDIQPKTH 527
Db 483 LRGIIRDLEATAASGHVEQLAVHLAQVDRPRELVEQRGRGFQNPQAGAEIVGGAQR 542
QY 528 QKPEARMPLSQCKGPDIFHRLGPLSVFSKAKNRWLVPVHLTRGEGGFLTLRGDSPVL 587
Db 543 QDQARLASPARQR-----FCHLAQRAVATAGDQRIAGRQLL-----DDAPGV 587
QY 588 IAAVPGSQAAGLKEGDIYVSVNGQPCRWHEAEVTELKAAAGEAGASLQVVS-----643
Db 588 --AVJPG-----YSHRQRPGLALRRDRGTNLFVARLLAQ 621
QY 644 -----LPSSRLPSLGDRRPVLIG 661
Db 622 DQONLVLVHRTPLVVALSLPIDRVSGTEVRPVGAG 658

RESULT 15
US-09-883-096-2
; Sequence 2, Application US/09883096
; Patent No. 6680369
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Craven, Andrew
; APPLICANT: Yu, Ming
; APPLICANT: Sakowicz, Roman
; APPLICANT: Patel, Umesh A.
; APPLICANT: Davies, Katherine A.
; TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE
; FILE REFERENCE: 020552-00141005
; CURRENT APPLICATION NUMBER: US/09/883,096
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 09/594,655
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 864
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Amino acid sequence encoded by human kinesin motor
; OTHER INFORMATION: protein gene HsKip3a (Figure 1).
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid
; OTHER INFORMATION: sequence of HsKip3a.
US-09-883-096-2
```

```
Query Match 3.4%; Score 128; DB 4; Length 864;
Best Local Similarity 20.1%; Pred. No. 0.0084;
Matches 129; Conservative 70; Mismatches 258; Indels 186; Gaps 24;

QY 211 PSRNESGLELIT-AYYNQLCFDARFLTPARSLGLFFHHYDSLTGVPAAQQALAFKGSV 269
Db 193 PASAEQLLEILTRGNRRNTQHTDANATSSRSHAIFQIFVKQDQRPVGLTQAVQVAKMSL 252
QY 270 LFNIGALHTQICARQDRSCTEGA--RRAMEAFORAAGAFSLLENFNSHAPSPDMSAASLC 327
Db 253 IDLAGSERASTHAKGEFLREGANINRSLALLINVLNADAKRKTHVPYRDSKLTREL- 311
QY 328 ALEQIMMAQAQECVFEGLSPSPASMAPQDCLAQLIAQEA-----366
Db 312 -LKDSLGNCRNTVMIAAIS-FSSLTYEDTYNTLVYADRAKEIRLSKSNVTSLDCHISQY 369
QY 367 -----AQVAA--EYELVHRTVAQPPVHDY--VPVSWTALVHVKAIFY-----404
Db 370 ATICOQLQAEVAALRKLQVYEGGQPPQDLPGSPKSGPPPEHLPSPLPPHPPSPQCT 429
```

```
QY 405 -----RSLAHYHVAMAL-----CDGSPATEGELPTHE---QVFLQPPTSSK 442
Db 430 PELPAGPRALQBEESLQMEAQVERAMEGNSSDQEQSPEDDEDEGFAEVPTQMPQNFTHAL 489
QY 443 PRGPVLPOELEERRQLGKAHLKRAILGQBEALRLHALCRVLRVDLIRA--VISQTLQRS 500
Db 490 PESPL--TLQPKPVVGHFSARELDGDRSKQLALKVLCVAQRQYSLLOANLLTPDMITE 547
QY 501 LAKYAEILDREDDFCEAAE-----APDIQPKTHQKPEARMPLSQCKGPDIFHRLG 550
Db 548 FETLQQLVCEEKIEPCAEALRTSGLARGAPLAQELCSESIIPVPSPLCPEPPG-----YTG 602
QY 551 PLSVFSKAKNRWLVPVHLTRGEGGFLTLRGDSPVLIAAVIPGSAAGLKEGDIYVS 610
Db 603 PVTRTMAR--RLSGPLH-----TL-----GIPPGNCTPA-----630
QY 611 VNGQPCRW-----WRHAEVVTTELKAAAGEAGASLQVVSLLPSRLPSLGDRRP-----657
Db 631 ---QSRWPMKRRRRPSALEADSPMAKRGYKROQSFCLRRGSLPDTQPSQGPSTP 687
QY 658 -----VLLGPRGLIIRSOREHGCKTP-----ASTWASPR 685
Db 688 KGERASSPCHSPRVCPATVIKSRVELGPSAM-----QNCSTPLALPTFDLNAITDLSSE 741
QY 686 P-----LLNWSRKAQCGTGGCPCQCAPVKPAPPSLKH 720
Db 742 PPSKPSFHECIGWDKIPQE--LSRLDQPFIPRAPVPLFTMKGP 782
```

Search completed: July 13, 2004, 12:08:40
Job time : 33 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2004, 19:58:24 ; Search time 637.213 Seconds
(without alignments)
16460.437 Million cell updates/sec

Title: US-10-697-266-1

Perfect score: 2469

Sequence: 1 tcgcggcccggtggtgcgg.....aaaaaaaaaaaaaaaaaaaaa 2469

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_29Jan04.*

1: geneseqn1980s.*

2: geneseqn1990s.*

3: geneseqn2000s.*

4: geneseqn2001as.*

5: geneseqn2001bs.*

6: geneseqn2002s.*

7: geneseqn2003as.*

8: geneseqn2003bs.*

9: geneseqn2003cs.*

10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query % Match	Score	Length	DB ID	Description
1	2469	100.0	2469	6	AAL48837 Human pro
2	2105	85.3	2837	7	ABV72511 Nucleotid
3	957	38.8	1013	5	AAS00831 Human cDN
4	955.4	38.7	1013	3	AAC77780 Human can
5	952.8	38.6	3213	7	ACC46698 Human dit
6	524.6	21.2	764	2	AZ15546 Human gen
7	496.4	20.1	2100	5	AAS15828 Human cDN
8	496.4	20.1	3526	5	AAS15827 Human ORF
9	494.8	20.0	2109	6	ABX97183 Human NOV
10	493.2	20.0	2061	6	ABX971856 Human GTP
11	493.2	20.0	3484	6	ABX971856 Human GTP
12	487.8	19.8	3019	4	AAF58361 Human cDN
13	463.6	18.8	2310	6	ABX97182 Human NOV
14	418.2	16.9	2757	4	AAC91349 Human pol
15	379.2	15.4	1671	6	ABK92255 Prostate
16	368	14.9	19025	6	AAL48838 Human pro
17	355	14.4	3335	5	AAS72556 DNA encod
18	301.4	12.2	2869	4	AAH75576 Human tra
19	235.4	9.5	966	5	AAS70406 DNA encod
20	219.4	8.9	331	3	AAC76200 Human ORF
21	206.6	8.4	2168	4	ABL14125 Drosophil
22	203.8	8.3	735	3	AAC75784 Human ORF
23	165	6.7	165	2	AAL14379 Human gen

24	150	6.1	933	5	AAS72551	Aas72551 DNA encod
25	143.8	5.8	1330	5	AAS70407	Aas70407 DNA encod
26	117.6	4.8	756	6	ABQ31046	Abq31046 Oligonucl
27	117.6	4.8	756	6	ABQ31047	Abq31047 Oligonucl
28	108.4	4.4	2925	7	ADA52888	Ada52888 Human cod
29	108.4	4.4	5200	6	AAD40740	Aad40740 Human kin
30	108.2	4.4	5234	3	AAC81224	Aac81224 Human his
31	103.6	4.2	599	4	AAL00243	Aal00243 Human rep
32	97	3.9	802	4	AAS03028	Aas03028 Human rep
33	94.6	3.8	439	8	ACH17102	Ach17102 Human adu
34	93.4	3.8	338	4	AAS57421	Aas57421 cDNA #97
35	87	3.5	479	6	ABL81142	Ab181142 Human ova
36	84	3.4	756	6	ABQ31049	Abq31049 Oligonucl
37	84	3.4	756	6	ABQ31048	Abq31048 Oligonucl
38	72.4	2.9	500	6	ABS71881	Abx71881 Human GTP
39	71	2.9	167	5	AAS15837	Aas15837 Human pro
40	71	2.9	389	5	AAS72552	Aas72552 DNA encod
41	69.4	2.8	167	6	ABS71866	Abx71866 Human GTP
42	68.2	2.8	789	5	AAS72550	Aas72550 DNA encod
43	68	2.8	2469	6	AAL48837	Aal48837 Human pro
44	68	2.8	2837	7	ABV72511	Abv72511 Nucleotid
45	66.2	2.7	492	3	AAC74860	Aac74860 Human ORF

ALIGNMENTS

RESULT 1
AAL48837
ID AAL48837 standard; cDNA; 2469 BP.
XX
AC AAL48837;
XX
DT 24-OCT-2002 (first entry)
XX
DE Human protein kinase N family kinase coding sequence.
XX
KW Human; protein kinase N; kinase; PKN; eye retinoblastoma; bocio tumour;
KW placenta choriocarcinoma; germ cell; leukocyte; uterus tumour; cancer;
KW pre-B cell acute lymphoblastic leukaemia; Wilm's tumour; enzyme;
KW brain anaplastic oligodendroma; uterus endometrial adenocarcinoma;
KW cystostatic; gene therapy; chromosome 8; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..102
FT CDS /*tag= a
FT 103..2274
FT 3'UTR /*tag= b
FT /*product= "kinase"
FT 2275..2469
FT /*tag= c
XX
EN WO200261062-A2.
XX
PD 08-AUG-2002.
XX
PF 29-JAN-2002; 2002WO-US002152.
XX
PR 01-FEB-2001; 2001US-00773371.
PR 07-MAY-2001; 2001US-00849334.
XX
PE (PEKE) PE CORP NY.
XX
PI Rusch D, Ketchum KA, Di Francesco V, Beasley EM;
XX
DR WPI; 2002-608516/65.
XX
PT P-PSDB; AAO18602.
XX
PT New human kinase peptide and nucleic acid molecule, useful for treating
PT disorders associated with abnormal expression of kinase protein, e.g.
PT retinoblastoma, Wilm's tumor, in drug screening assays and

PT pharmacogenomic analysis.

XX PS Claim 4; Fig 1; 76pp; English.

XX CC The present invention provides the protein and coding sequences of a novel human protein kinase N family protein. The sequences can be used in the treatment of disorders associated with the absence of, inappropriate, or unwanted expression of the protein, e.g. eye retinoblastoma, Wilms' tumour, placenta choriocarcinomas, bocio tumours, pre-B cell acute lymphoblastic leukaemias, uterine tumours, brain anaplastic oligodendromas and uterine endometrial adenocarcinomas. The present sequence is the cDNA of the invention. The gene is found on human chromosome 8

XX SQ Sequence 2469 BP; 461 A; 812 C; 821 G; 375 T; 0 U; 0 Other;

Query Match 100.0%; Score 2469; DB 6; Length 2469;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2469; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TCGCGGCCCAAGTGTGTCGGGCGGCGCCCTAGCCCGGCTGCGGAGCGCTGCGCGAGCGCGG	60
DB	1	TCGCGGCCCAAGTGTGTCGGGCGGCGCCCTAGCCCGGCTGCGGAGCGCTGCGCGAGCGCGG	60
QY	61	GCTGGCTGACCCCGAGGAGACCCCGAGCGAGCGGCTGCGGCGATGATCTTGGAGGAGG	120
DB	61	GCTGGCTGACCCCGAGGAGACCCCGAGCGAGCGGCTGCGGCGATGATCTTGGAGGAGG	120
QY	121	CCGAGCGGCGCGCGCGCGAGGAGAGCCCGCGGCTGCAGATATCTAGGAGAAACCC	180
DB	121	CCGAGCGGCGCGCGCGCGAGGAGAGCCCGCGGCTGCAGATATCTAGGAGAAACCC	180
QY	181	AGGAAACACGTGTAGCTCTTTACGGGGAAGACGGGAGCGCTGAGAGACGTGTGGG	240
DB	181	AGGAAACACGTGTAGCTCTTTACGGGGAAGACGGGAGCGCTGAGAGACGTGTGGG	240
QY	241	TGAGAGGGGTGTGCGGGTCCACAGAGGGGAAGACCCAGTGGCTGTGCACTGGGCCCATG	300
DB	241	TGAGAGGGGTGTGCGGGTCCACAGAGGGGAAGACCCAGTGGCTGTGCACTGGGCCCATG	300
QY	301	AATCCGCGAGTTCAATGAGTGGGTGTGATCTCCCTGACGAGATCCAGTGGCGGACGCTG	360
DB	301	AATCCGCGAGTTCAATGAGTGGGTGTGATCTCCCTGACGAGATCCAGTGGCGGACGCTG	360
QY	361	CAGAGCCGAGGGCCAGATTCACACAGATTCACAGAGGAGCTGCAGATGCGGAGCGG	420
DB	361	CAGAGCCGAGGGCCAGATTCACACAGATTCACAGAGGAGCTGCAGATGCGGAGCGG	420
QY	421	GCTGAGAACCTCTACAGAGCCACAGCAACACCGGGTGAAGAGACGGTTCGCCCTGGAG	480
DB	421	GCTGAGAACCTCTACAGAGCCACAGCAACACCGGGTGAAGAGACGGTTCGCCCTGGAG	480
QY	481	CTGAGCTACGTCAAATCCGAGCTGCTGAGGAGGAGCTGAGGAGGAGCTCAGCGGT	540
DB	481	CTGAGCTACGTCAAATCCGAGCTGCTGAGGAGGAGCTGAGGAGGAGCTCAGCGGT	540
QY	541	GGGTGACCTGCGCGGCTAGGAGGAGGAGCTGCTCACTGTCCTCATGATCCCGCTGGG	600
DB	541	GGGTGACCTGCGCGGCTAGGAGGAGGAGCTGCTCACTGTCCTCATGATCCCGCTGGG	600
QY	601	CTGAAGGAGACCAAGGAGCTGGACTGTCTACACCGCTGAAGGAGCTGATCTAGTGAC	660
DB	601	CTGAAGGAGACCAAGGAGCTGGACTGTCTACACCGCTGAAGGAGCTGATCTAGTGAC	660
QY	661	TTTGGAGAGAGCGGCGCTCTACAGGACAAATCAGGAGCTGAGGCGCTCGGGCAG	720
DB	661	TTTGGAGAGAGCGGCGCTCTACAGGACAAATCAGGAGCTGAGGCGCTCGGGCAG	720
QY	721	GCCATGCGGACCCCGAGCGGAATGAGTGGGCTGGAGCTGCTCAGAGCTATTACAA	780
DB	721	GCCATGCGGACCCCGAGCGGAATGAGTGGGCTGGAGCTGCTCAGAGCTATTACAA	780
QY	781	CAGCTGTCTTCTTGATGCGCGCTTCTCACCCTGCGAGGAGCTCGGGCTCTTCTTC	840

DB	781	CAGCTGTGCTTCTTGATGCGCGCTTCTCACCCTGCGAGGAGCTCGGGCTCTTCTTC	840
QY	841	CACTGGTAGAGTCCGCTTACTGGGGTCCCGGCCACGACGCTGCGCTTCCGAGAAG	900
DB	841	CACTGGTAGAGTCCGCTTACTGGGGTCCCGGCCACGACGCTGCGCTTCCGAGAAG	900
QY	901	GGCAGGGTCTCTTCAACATCGGTGCCCTCCACACGACAGATTGGGGCGGCCAGGCCG	960
DB	901	GGCAGGGTCTCTTCAACATCGGTGCCCTCCACACGACAGATTGGGGCGGCCAGGCCG	960
QY	961	TCTGCAACGAGGGTCCCGCGCTATGAGAGCTTCCAGAGGCGCTGGGGCGCTTC	1020
DB	961	TCTGCAACGAGGGTCCCGCGCTATGAGAGCTTCCAGAGGCGCTGGGGCGCTTC	1020
QY	1021	AGCCTCTGAGGAGAACTTCTCCATGCGCGGAGCCACAGATAGAGCGCTGCTCCCTC	1080
DB	1021	AGCCTCTGAGGAGAACTTCTCCATGCGCGGAGCCACAGATAGAGCGCTGCTCCCTC	1080
QY	1081	TGCGCACTGAGCAGCTCATGATGGCCAGGCCAGGAATGTGTGTTGAGGCGCTCTCA	1140
DB	1081	TGCGCACTGAGCAGCTCATGATGGCCAGGCCAGGAATGTGTGTTGAGGCGCTCTCA	1140
QY	1141	CCACTGCTCCATGCGCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1200
DB	1141	CCACTGCTCCATGCGCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1200
QY	1201	GCCAGGTGCGACGAGTACAGGCTAGTGCACCGGACCATGCGCCACGACCGCTCCAC	1260
DB	1201	GCCAGGTGCGACGAGTACAGGCTAGTGCACCGGACCATGCGCCACGACCGCTCCAC	1260
QY	1261	GACTAGTGCCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1320
DB	1261	GACTAGTGCCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1320
QY	1321	CTGGCCCACTACACAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1380
DB	1321	CTGGCCCACTACACAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1380
QY	1381	CTCCCAACGACAGAGGCTTCTTCTGACGCGCCCGACCTCTCTAAGCGCCCGAGGCGCT	1440
DB	1381	CTCCCAACGACAGAGGCTTCTTCTGACGCGCCCGACCTCTCTAAGCGCCCGAGGCGCT	1440
QY	1441	GTCTGCTGCGCAGAGCTGAGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1500
DB	1441	GTCTGCTGCGCAGAGCTGAGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1500
QY	1501	ATCTGCGGCGAGGAGGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG	1560
DB	1501	ATCTGCGGCGAGGAGGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG	1560
QY	1561	GACTGCTTCCGGGCTGTGATCTCCAGACGCTGCGAGCTCTCTAAGCGCCCGAGGCGCT	1620
DB	1561	GACTGCTTCCGGGCTGTGATCTCCAGACGCTGCGAGCTCTCTAAGCGCCCGAGGCGCT	1620
QY	1621	CTGACCGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1680
DB	1621	CTGACCGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1680
QY	1681	CACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1740
DB	1681	CACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1740
QY	1741	CATCGGCTGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1800
DB	1741	CATCGGCTGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1800
QY	1801	GTCACCTGACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1860
DB	1801	GTCACCTGACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1860
QY	1861	CTCATCGCTGCGCTCATTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1920
DB	1861	CTCATCGCTGCGCTCATTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1920

XX	12-MAR-1999;	99US-0124270P.	
PR	(HUMA-) HUMAN GENOME SCI INC.		
XX	Rosen CA, Ruben SM;		
XX	WPI; 2000-587533/55.		
DR	P-PSDB; AAB43571.		
XX	Novel isolated nucleic acids comprising sequences encoding peptides		
PT	useful for treating or diagnosing e.g. cancer.		
PT	Claim 1; Page 752-753; 2352pp; English.		
XX	AAC77607 to AAB43398 encode the human cancer associated proteins given in		
XX	AAB43398 to AAB43571. The proteins can have activities based on the		
CC	tissues and cells the genes are expressed in. Example of activities		
CC	include: cytostatic; proliferative; immunomodulator;		
CC	antidiabetic; antiasthmatic; antirheumatic; antiarthritic;		
CC	antiinflammatory; antihypertensive; antiallergic; antibacterial;		
CC	dermatological; neuroprotective; cardiatic; thrombolytic; coagulant;		
CC	nootropic; vasotropic; antipsoratic and antiangiogenic. The		
CC	polynucleotides and polypeptides can be used for preventing, treating or		
CC	ameliorating medical conditions and diagnosing pathological conditions.		
CC	Polynucleotides, polypeptides, antibodies, agonists and antagonists from		
CC	the present invention may be used to treat immune disorders by activating		
CC	or inhibiting the proliferation, differentiation or mobilisation of		
CC	immune cells, to treat disorders of haematopoietic cells, autoimmune		
CC	disorders, allergic reactions, graft versus host disease and organ		
CC	rejection, modulate haemostatic or thrombolytic activity, modulate		
CC	inflammation, cancers, cardiovascular disorders, neurological disease and		
CC	bacterial or viral infections. The peptides, nucleotides, antibodies,		
CC	agonists and antagonists may be also be used in drug screens. AAC78449 to		
CC	AAC78457 and AAB44240 represent sequences used in the exemplification of		
CC	the present invention		
XX	Sequence 1013 BP; 197 A; 328 C; 337 G; 151 T; 0 U; 0 Other;		
SQ	Query Match 38.7%; Score 955.4; DB 3; Length 1013;		
	Best Local Similarity 99.2%; Pred. No. 2.9e-163;		
	Matches 981; Conservative 0; Mismatches 6; Indels 2; Gaps 2;		
Qy	1477 GCGAGGACACCTGAAGCGTGCATCTCTGGGCGAGGAGGCGCTGCGGCTGACGCC 1536		
Db	27 GCGAAGGACACCTGAAGCGTGCATCTCTGGGCGAGGAGGCGCTGCGGCTGACGCC 86		
Qy	1537 CTGTGCGCGTCTCTGCGGAGTGACCTGTCTTGGGCTGTGTATCTCCAGACGCTGCAG 1596		
Db	87 CTGTGCGCGTCTCTGCGGAGTGACCTGTCTTGGGCTGTGTATCTCCAGACGCTGCAG 146		
Qy	1597 CGCTCACTGGCCAAAGTAGCGAGCTCGACCGTGAAGATGACTTCTGTGAGGCTGCCGAG 1656		
Db	147 CGCTCACTGGCCAAAGTAGCGAGCTCGACCGTGAAGATGACTTCTGTGAGGCTGCCGAG 206		
Qy	1657 GCGCGGACATCCAGCTTAGACCCACAGAGCCAGAGGCGAGATGCGACGCTGTGCC 1716		
Db	207 GCGCGGACATCCAGCTTAGACCCACAGAGCCAGAGGCGAGATGCGACGCTGTGCC 266		
Qy	1717 CAGGGGAAGGGGCTGTGACATCTTCCATCGGCTGGGGGCCCTGTCTGTGTCTTCAGGCCAAG 1776		
Db	267 CAGGGGAAGGGGCTGTGACATCTTCCATCGGCTGGGGGCCCTGTCTGTGTCTTCAGGCCAAG 326		
Qy	1777 AACCGGTGGCGGTGGGGCCCGTCCACTGTACCGAGAGAGGGCGGCTTTGGGCTC 1836		
Db	327 AACCGGTGGCGGTGGGGCCCGTCCACTGTACCGAGAGAGGGCGGCTTTGGGCTC 386		
Qy	1837 ACGCTTCGGGAGACTCCCTGTCTCATCGCTGCGCTCATTCCAGGGAGCCAGGCCCGC 1896		
Db	387 ACGCTTCGGGAGACTCCCTGTCTCATCGCTGCGCTCATTCCAGGGAGCCAGGCCCGC 446		
Qy	1897 GCGGCTGCGCTGAAGGAGGGCGACTACATGTGTGAGTGAATGGGACGCGATGCGAGTGG 1956		
Db			

Db	447 GCGGCTGGCCTGAAGGAGGGCGACTACATGTGTGTGAGTGAATGGGCGAGCCATGCGAGGTGG 506
Qy	1957 TGGAGACACGCGAGGTGTGACGAGCTGAAGGCTGCGGAGAGAGCGGCGCGCCAGCCCTG 2016
Db	507 TGGAGACACGCGAGGTGTGACGAGCTGAAGGCTGCGGAGAGAGCGGCGCGCCAGCCCTG 566
Qy	2017 CAGGTGGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACCGCGGCGCCCTG 2076
Db	567 CAGGTGGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACCGCGGCGCCCTG 626
Qy	2077 CTGCTGGGGCCCGAGGGGGCTTCTAAGGACGAGAGGAGCATGGTTGCAAGACCCCGGCA 2136
Db	627 CTGCTGGGGCCCGAGGGGGCTTCTAAGGACGAGAGGAGCATGGTTGCAAGACCCCGGCA 686
Qy	2137 TCCACGTGGGGCGAGTCCCGGCGCCCTCTCAACTGAGAGCCGAAAGGCCAGAGGCGCAAG 2196
Db	687 TCCACGTGGGGCGAGTCCCGGCGCCCTCTCTAAGTGAAGCGGCGGCGGCGGCGCAAG 746
Qy	2197 ACTGGAAGGCTGCCCCCAGAGCCCTGTGTGCCAGTGAAGCCAGCTCCGCGCTCATCTTGAAG 2256
Db	747 ACTGGAAGGCTG-CCCCAGAGCCCTGTGTGCCAGTGAAGCCAGCTCCGCGCTCATCTTGAAG 805
Qy	2257 CACCCAGGCTGGCCGTGAGGCGGCGAGTCCCTGACGCGCCCTCAGCCCTGCTCCAGCTGG 2316
Db	806 CACCCAGGCTGGCCGTGAGGCGGCGAGTCCCTGACG-CTTACGCGCTGCTCCAGCTGG 864
Qy	2317 CAGCAAGCAGCCAGAGTGCCTCCCGACCCAGAGGACCTCCCGGCAATGCTGTCGCGCC 2376
Db	865 CAGCAAGCAGCCAGAGTGCCTCCCGACCCAGAGGACCTCCCGGCAATGCTGTCGCGCC 924
Qy	2377 TCAATGCTGAGGCTGCTCGGGCACTGCTGCGGCACTTAAAGACTGGTCAAGCTGTCTG 2436
Db	925 TCAATGCTGAGGCTGCTCGGGCACTGCTGCGGCACTTAAAGACTGGTCAAGCTGTGAAAA 984
Qy	2437 AAAAAA AAAAAAAAAAAAAAAAAAAAAA 2465
Db	985 AAAAAA AAAAAAAAAAAAAAAAAAAAAA 1013

RESULT 5	
ACC46698	
ID	ACC46698 standard; cDNA; 3213 BP.
XX	
AC	ACC46698;
XX	
DT	02-JUN-2003 (first entry)
XX	
DE	Human dithp biochemical pathway protein-encoding cDNA.
XX	
KW	Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
KW	cancer; cell proliferative disorder; autoimmune disorder;
KW	inflammatory disorder; infection; hormonal disorder; metabolic disorder;
KW	neurological disorder; gastrointestinal disorder; transport disorder;
KW	connective tissue disorder; drug screening; proteome analysis;
KW	gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
KW	disease model; toxicological testing; transcript imaging;
XX	biochemical pathway; gene; ss.
OS	Homo sapiens.
XX	
FN	WO200297031-A2.
XX	
PD	05-DEC-2002.
XX	
PF	27-MAR-2002; 2002WO-US010056.
XX	
PR	28-MAR-2001; 2001US-0279619P.
PR	29-MAR-2001; 2001US-0280067P.
PR	16-MAY-2001; 2001US-0280068P.
PR	17-MAY-2001; 2001US-0291280P.
PR	17-MAY-2001; 2001US-0291829P.
PR	19-JUN-2001; 2001US-0299428P.

```
PR 20-JUN-2001; 2001US-0299776P.
PR 20-JUN-2001; 2001US-0300001P.
XX (INCY-) INCYTE GENOMICS INC.
XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
PI Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX WPI; 2003-129518/12.
DR P-FSDB; ABR41761.
XX
XX Novel human diagnostic and therapeutic polypeptide useful for identifying
PT test compound which specifically binds to a polypeptide encoded by human
PT diagnostic and therapeutic polynucleotide, and to induce antibodies.
XX
XX Claim 2; SEQ ID NO 619; 591pp; English.
XX
XX The invention relates to novel human diagnostic and therapeutic
CC polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded
CC proteins (DITHP; ABR41136-ABR41812). The invention also relates to
CC polynucleotide sequences at least 90% identical to the dithp cDNA
CC sequences of the invention; recombinant vectors, host cells and
CC transgenic organisms comprising a dithp nucleic acid sequence; the
CC recombinant production of DITHP proteins; antibodies specific for DITHP
CC proteins; microarrays comprising dithp nucleic acid sequences; methods of
CC detecting dithp nucleotide and protein sequences; methods of screening
CC for compounds which specifically bind a DITHP protein; and methods of
CC assessing the toxicity of test compounds using a dithp hybridisation
CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the
CC diagnosis of a wide variety of conditions including cancer and other cell
CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,
CC viral, fungal or parasitic infections; hormonal disorders; metabolic
CC disorders; neurological disorders; gastrointestinal disorders; transport
CC disorders; and connective tissue disorders. They may also be used to
CC screen for modulators of protein activity or gene expression. DITHP
CC proteins can additionally be used in analysis of the proteome of a tissue
CC or cell type and to induce antibodies. The dithp nucleic acids are
CC additionally useful in somatic or germline gene therapy of the disorders
CC mentioned above, as a source of antisense sequences, as a source of
CC probes and primers, in genotyping and identification of individuals, in
CC the generation of transgenic animal models of human disease or knock in
CC humanised animals, in toxicological testing, and in transcript imaging.
CC The present sequence represents a dithp cDNA encoding a DITHP protein
CC which is involved in a biochemical pathway. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 3213 BP; 538 A; 1114 C; 1029 G; 532 T; 0 U; 0 Other;
Query Match 38.6%; Score 952.8; DB 7; Length 3213;
Best Local Similarity 93.2%; Pred. No. 9.8e-163;
Matches 1018; Conservative 0; Mismatches 72; Indels 2; Gaps 2;
QY 1346 CCCTCTGCGAGCGGTCTCCAGCGACCGAGGGAGAGCTCCCAACGACGAGCAGGTCTTCC 1405
DB 712 CCCGTGCGCGGCACCCCAAGAAAGTGGTGTGATGAGCCCAACAGCCCTGGCGTTGC 771
QY 1406 TGCAGCCCCCACCCTCTTAAGCCCCGAGGCCCTGTGCTGCCGAGGAGCTGGAGGAGC 1465
DB 772 CCACCTCTTCTGCCACGTCCAGGGGCCACGCGGCCACATGGTGTGACATCCCGATGC 831
QY 1466 GCAGGACGTTGGCAAGGACACCTGAAGCGTCCATCTCTGGGGCAGGAGGGCGTGC 1525
DB 832 CCCGGTGC-AGGCAAGGACACCTGAAGCGTCCATCTCTGGGGCAGGAGGGCGTGC 890
QY 1526 GGCTGCAGCCCTGTGCGCGCTCTTCCGCGAGGTGGACCTGCTTCCGGCTGTGATCTCCC 1585
DB 891 GGCTGCAGCCCTGTGCGCGCTCTTCCGCGAGGTGGACCTGCTTCCGGCTGTGATCTCCC 950
1586 AGACGCTCAGCGCTCACTGCGCAAGTATGCGGAGCTCGACCGTGAGGATGACTTCTGTG 1645
951 AGACGCTCAGCGCTCACTGCGCAAGTATGCGGAGCTCGACCGTGAGGATGACTTCTGTG 1010
1646 AGGCTGCGGAGGCCCGGACATCCAGCTTAAGCCACCAAGCCAGAGGCCAGGATGC 1705
1011 AGGCTGCGGAGGCCCGGACATCCAGCTTAAGCCACCAAGCCAGAGGCCAGGATGC 1070
1706 CACGCTGTCCACAGGGGAAGGGGCTGACATCTTCCATCGGCTGGGGCCCTCTGTGTGT 1765
1071 CACGCTGTCCACAGGGGAAGGGGCTGACATCTTCCATCGGCTGGGGCCCTCTGTGTGT 1130
1766 TCTCAGCCAAAGACCGGTGGCGGCTGTGGGGCCCGTCCACTGACCCAGAGAGGGCG 1825
1131 TCTCAGCCAAAGACCGGTGGCGGCTGTGGGGCCCGTCCACTGACCCAGAGAGGGCG 1190
1826 GCTTTGGGCTCAGCTTCGGGGAGAGCTGCGCTGTCTCTCATGCTGCGCTCATTCAGGGA 1885
1191 GCTTTGGGCTCAGCTTCGGGGAGAGCTGCGCTGTCTCTCATGCTGCGCTCATTCAGGGA 1250
1886 GCCAGGCCGCGCGGCTGCGCTGAAGAGGGGCGACTACATTTGTGTGATGAGTGGGCGC 1945
1251 GCCAGGCCGCGCGGCTGCGCTGAAGAGGGGCGACTACATTTGTGTGATGAGTGGGCGC 1310
1946 CATGACGCTGTGGAGACACCGCGAGGTGGTGAAGGAGCTGAAGGCTCGCGGAGAGGGCG 2005
1311 CATGACGCTGTGGAGACACCGCGAGGTGGTGAAGGAGCTGAAGGCTCGCGGAGAGGGCG 1370
2006 GCGCCAGCTCAGGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2065
1371 GCGCCAGCTCAGGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1430
2066 GCGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2125
1431 GCGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1490
2126 AGACCCCGGATCAACGCTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2185
1491 AGACCCCGGATCAACGCTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1550
2186 AGCAGGGCAAGACTGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2245
1551 AGCAGGGCAAGACTGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1610
2246 CATCTTTGAAGCACCCAGGCTGGCGCTGAGGGCCAGGATCCTGACGCGCTCAGCCCTG 2305
1611 CATCTTTGAAGCACCCAGGCTGGCGCTGAGGGCCAGGATCCTGACG-CTCAGCCCTG 1669
2306 GCTCCAGCTGGCAGCAAGCAGCAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2365
1670 GCTCCAGCTGGCAGCAAGCAGCAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1729
2366 CTTGTCGCGCTCATGCTGGAGGCTGCTGCTGCGGACCTGCTGCTGCTGCTGCTGCTGCTG 2425
1730 CTTGTCGCGCTCATGCTGGAGGCTGCTGCTGCGGACCTGCTGCTGCTGCTGCTGCTGCTG 1789
2426 AGACCTGTCTGA 2437
1790 AGACCTGTCTGA 1801
RESULT 6
AAZ15546
ID AAZ15546 standard; cDNA; 764 BP.
XX
AC AAZ15546;
XX
DT 12-OCT-1999 (first entry)
XX
DE Human gene expression product cDNA sequence SEQ ID NO:3015.
XX
KW Human; gene; gene expression product; diagnosis; therapy; probe;
KW detection; mapping; tissue typing; profiling; forensic; cancer;
```

KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX Homo sapiens.
OS
PN WO9938972-A2.
XX
PD 05-AUG-1999.
XX
PF 28-JAN-1999; 99WO-US001619.
XX
PR 28-JAN-1998; 98US-0072910P.
PR 24-FEB-1998; 98US-0075954P.
PR 31-MAR-1998; 98US-0080114P.
PR 03-APR-1998; 98US-0080515P.
PR 03-APR-1998; 98US-0080666P.
PR 21-OCT-1998; 98US-0105234P.
PR 28-OCT-1998; 98US-0105877P.
XX
PA (CHIR) CHIRON CORP.
PA (HYSR-) HYSRQ INC.
XX
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Glese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
PI Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
PI Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;
XX WPI; 1999-494092/41.
XX
XX Novel human genes and their expression products which are differentially
PT expressed in different cell types.
PT
XX
PS Claim 1; Page 1452; 2479pp; English.
XX
CC The present invention describes a library of human polynucleotides
CC comprising the sequences given in AAZ12532 to AAZ17779. Also described is
CC a method of detecting differentially expressed genes correlated with the
CC cancerous state of a mammalian cell, comprising detecting at least one
CC differentially expressed gene product in a test sample from a cell
CC suspected of being cancerous, where the gene product is encoded by one of
CC the 5248 polynucleotide sequences given in AAZ12532 to AAZ17779. The
CC polynucleotides can be used as a source of primers and probes, which can
CC be used for a variety of purpose, e.g. detection of expression levels,
CC mapping, tissue typing or profiling, forensics, genetic analysis and
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
CC can be used for raising antibodies for experimental, diagnostic and
CC therapeutic purposes. The polynucleotides may also be used to construct
CC arrays for diagnostics (which may be used to determine function of an
CC encoded protein); and to detect differences in expression levels between
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC identify a genetic predisposition or susceptibility to a disease such as
CC cancer). The polynucleotides of the invention are especially used in the
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC and lung cancer. The polynucleotides can also be used to screen for
CC peptide analogues and antagonists
XX
SQ Sequence 764 BP; 157 A; 224 C; 224 G; 137 T; 0 U; 22 Other;

Query Match 21.2%; Score 524.6; DB 2; Length 764;
Best Local Similarity 92.6%; Pred. No. 1.7e-85;
Matches 601; Conservative 0; Mismatches 39; Indels 9; Gaps 5;

QY 1821 GGGCGGCTTGGGCTCACGCTTCGGGGAGACTCGCTGTCTCATGCTGCCGTCAATTC 1880
DB |||||
51 GGGCGGCTTGGGCTCACGCTTCGGGGAGACTCGCTGTCTCATGCTGCCGTCAATTC 110
QY 1881 AGGAGCCAGCCGCGCGCGCTGGCTTGAAGAGGGGAGCTACATTGTGTCAGTGAATGG 1940
DB |||||
111 AGGAGGACAGCCGCGCGCGCTGGCTTGAAGAGGGGAGCTACATTGTGTCAGTGAATGG 170
QY 1941 GCAGCCATGAGTGTGGAGACACACGCGAGGTGGTACCGAGCTGAAGCTCGGGAGA 2000
DB |||||
171 GCAGCCATGAGTGTGGAGACACACGCGAGGTGGTGGAGAGCTGAGAGCTCGGGAGA 230

QY 2001 GGGCGGCGCAGCTGCAGTGGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGG 2060
DB |||||
231 GGGCGGCGCAGCTGCAGTGGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGG 290
QY 2061 GGACCGCGCGCGCTCTGCTGGGCCCCCAGGGGGCTTCTAAGAGGCAGAGGAGCATGG 2120
DB |||||
291 GGACCGCGCGCGCTCTGCTGGGCCCCCAGGGGGCTTCTAAGAGGCAGAGGAGCATGG 350
QY 2121 TTGCAAGACCCCGGCATCCAGTGGGCGGAGTCCCGGGCCCTCTCAACTGGAGCCGAAA 2180
DB |||||
351 TTGCAAGACCCCGGCATCCAGTGGGCGGAGTCCCGGGCCCTCTCAACTGGAGCCGAAA 410
QY 2181 GGCCAGCAGGGCAAGACTGGAGGCTGCCCCAGCCCTGTGCCCACTGAAGCCAGCTCC 2240
DB |||||
411 GGCCCAACAGGGCAAGACTGGAGGCTGCCCCAGCCCTGTGCCCACTGAAGCCAGCTCC 470
QY 2241 GCCTCATCTTTGAAGACACCCAGGGTGGCGGTGAGGGCCAGAGATCCCTGACGCCCTCAG 2300
DB |||||
471 G-CCTCATCTTTGAAGACACCCAGGGTGGCGGTGAGGGCCAGAGATCCCTGACGCC--TCA 527
QY 2301 CCTGGCTCCAGCTGGCAGCAGCAGCAGCATGCCCTCCCAACCCAGAGGACCTCGGG 2360
DB |||||
528 CCTGGCTCCCACTGGCAGCAGCAGCAGCATGCCCTTCCCAACCCAGAGGACCTTCNGG 587
QY 2361 CAATGCC-TGTCGCCGCTCATGCT--GGAGGCTGCCCTCGGGCACCTGCC---TSCCCATT 2414
DB |||||
588 CAATGCTTGTNCCGCTTATGCTTGAAGCTTGGCTGGGCGACCTTGCCTTGNCCATTT 647
QY 2415 AAGACTGGTCAGACCTGTCTGAAAAAATAAAAAAAAAAAAAAAAAAAAAA 2463
DB 648 AAGACTGGTCANAACTGAAAAAATAAAAAAAAAAAAAAAAAAACTTCGAGAAA 696

RESULT 7
AAS15828
ID AAS15828 standard; cDNA; 2100 BP.
XX
AC AAS15828;
XX
DT 16-JAN-2002 (first entry)
XX
DE Human cDNA encoding prostate specific protein PSL22.
XX
KW Human; prostate specific protein; PSL22; prostate cancer;
KW benign prostatic hyperplasia; acute prostatitis; testicular cancer;
KW cryptorchidism; testicular disorder; proliferative disorder; lymphoma;
KW leukaemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;
KW liver cancer; lung cancer; cytostatic; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2061
FT /tag= a
FT /product= "PSL22"
XX
XX WO200172962-A2.
XX
XX 04-OCT-2001.
XX
XX 23-MAR-2001; 2001WO-US009410.
XX
XX 24-MAR-2000; 2000US-0191929P.
XX
XX (SAAT/) SAATCIOGLU F.
XX
XX Saatcioglu F;
XX
XX WPI; 2001-662926/76.
XX
XX P-PSDB; AAU10192.
XX
XX New polynucleotide for the diagnosis, prevention and treatment for
PT prostate and testis disorders, particularly prostate cancer, comprises

PT prostate-specific or testis-specific nucleic acids.

XX Claim 4; Fig 14B; 114pp; English.

XX The invention relates to substantially pure prostate-specific or testis-specific polypeptides and the nucleic acids encoding them. Also included are vectors and host cells expressing the proteins, a transgenic animal expressing the protein, antibodies against the proteins, probes for detecting the nucleic acids, antisense molecules for the nucleic acids and methods of isolating modulators of the proteins. Compounds that modulate the prostate specific or testis specific polypeptide are useful to diagnose, prevent or treat disorders of the testis or prostate particularly prostate cancer, benign prostatic hyperplasia, acute prostatitis, testicular cancer, cryptorchidism, undescended, retractile, ascending or vanished testis. Other proliferative disorders for which the modulators may be used include lymphoma, leukaemia, melanoma, ovarian cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer. The present sequence encodes a prostate specific protein, PSU22

XX Sequence 2100 BP; 516 A; 569 C; 584 G; 431 T; 0 U; 0 Other;

Query Match 20.1%; Score 496.4; DB 5; Length 2100;

Best Local Similarity 56.8%; Pred. No. 2.3e-80;

Matches 977; Conservative 0; Mismatches 731; Indels 12; Gaps 3;

QY 321 GGGCTGTGACTCCCTGACGACATCCAGTGGCGGCGAGCTGACGCGCGCCAGAT 380

DB 69 GGGCTGTAAATCCCTTGCACAAACCGCGCGAGTAAATTGCAGATCAAGAGCTGCTT 128

QY 381 TCACCAGCAGATTGACAGGAGCTCAGATGCGGAGCGGCGTGTAGAACCTTACAGAGC 440

DB 129 GAATCAGCAGATCTTGAAGCGCTGCGATGAGGACCGGAGCGGAACCTTCTGAAAGT 188

QY 441 CACCAGCAACACCGGCTGAGAGACGCTGCCCTGGAGCTGAGCTCACTCACTCCAA 500

DB 189 GCCCAAACTCAAGGTGCGGACAAAGTGGCGCTGGAGCTGAGCTTCTGCTCACTCAGA 248

QY 501 CTTGAGCTGTGAAAGGAGCTGAGAGCTCAGCGGTGGCGTGGACCTTGGCGCGCA 560

DB 249 CTTGAGATGCTCAAGGAAGAGCTGAGGCGGTGAACATCTCGGTGGCGCTTATCAGAA 308

QY 561 TGGAGCGAAGCTGTCACTGTCCCATGATCCCTTGGCGCTGAGGAGCAAGAGAGCT 620

DB 309 CACAGAGGAGGATTTAGATTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 368

QY 621 GGACTGTGCTACACCGCTGAAGGAGCTGATCTCAGTGCATTTTGGAGAGGAGCGGCTC 680

DB 369 CGACTTTCAGTGTCTCTCAGGATTTTATCTGGAACATTTACAGTGAAGTGGCTATTT 428

QY 681 CTACAGGCGAAGATCAAGGAGCTGAGGCGCTGCGGAGGCGCATGCGACCCCGCGCG 740

DB 429 ATATGAAGATGAATTTGAGATCTTATGGATCTGAGACAAGCTTGTGCGAGCGCTAGCG 488

QY 741 GAATGAGTGGCGCTGGAGCTGCTCAGAGCTATTACAAACAGCTGTGCTTCTGATGC 800

DB 489 GGAAGAGCGGCGGTGGAACCTGCTGATGACATCTTCACTCAGCTGGGCTTGTGCGAG 548

QY 801 CGCGTCTCTCAACCGCTGCGAGGAGCTCGGGCTCTTCTTCCACTGTGTACGACTGCTTAC 860

DB 549 TCGATTCTTCCGCGCCACAGCGCAGATGGACTCTGTTTACCTGTTATGATCTCTCAC 608

QY 861 TGGGCTCCCGGCCAGCAGCGTGGCGCTTCCGAGAGGCGAGCGTCTTCTTCAACAT 920

DB 609 TGGGCTCCCGGTGAGCAGCAGAACCTTCTGCTGAGAGGCGCAGTGTCTCTTCAACAC 668

QY 921 CGGTGCTTCCACAGCAGATTTGGGCGCGCAGAGCGCTCTCTGCAACGAGGTGGCCG 980

DB 669 TGGGCGCTCTACACCCAGATTTGGGACCGCGGTGCGATCGCGAGCGAGGCTGGGCTGGA 728

QY 981 CCGCGCTATGAGGCGCTTCCAGAGGCGCGCTGGGCGCTTTCAGCGCTCTGAGGAGAACTT 1040

DB 729 GAGTGCCATAGATGCTTTCAGAGAGCGCGAGGGGTTTAAATTAACCTTGAAGACACATT 788

QY 1041 CTCCCATGCGCGAGCCCGACATGAGCGCTGCTCCTCTGCGCACTGGAGCAGCTCAT 1100

DB 789 TACCCATACTCCAACTTACGACATGAGCCCTGCGCATGCTCAGCGTGTCTCAAAATGAT 848

QY 1101 GATGGCCAGGCGCCAGGAATGTGTTTGGGGCCCTCTCACACCTGCTCCATGCGCC 1160

DB 849 GCTTGCACAAGCCCAAGAAAGCGTGTGAGAAAATCAGCCTTCTCTG-----GGATCGG 902

QY 1161 CCAAGACTGCTGGCGCCAGCTGCGCTGCGCAGAGCGCGCCAGAGTGGAGCGAGTA 1220

DB 903 GAATGAATCTTCACTGTGTGTAAGGTGCTCAGAGGCTGCTAAGGTGGAGAGGTCTA 962

QY 1221 CAGGCTAGTGACCGGACCATGCGCCAGCCACCCCTCCACGACTAGTGCCTGTCTCTTG 1280

DB 963 CCAACAGCTACAGGCGCATGAGCAGCGCGGTGAAGAGAACATCCCTACTCTCTCTG 1022

QY 1281 GACTGCCCTGGTGATGTCAAGCGCGAGTACTTCCGCTCCCTGGGCCCACTACAGTACG 1340

DB 1023 GGGCAGCTTAGCCTGCGTGAAGGCCCACTACGCGGCGCTTGGCCCACTTCTTCACTGC 1082

QY 1341 CATGGCCCTCTGCGAGCGCTCCCGACGACCGAGGAGAGCTCCCGACGCGAGCAGGT 1400

DB 1083 CATCTCTCTATCGACACACAGGTGAACCGAGGCAAGGCTGAGACCCAGAGAGAGTG 1142

QY 1401 CTTCTCTGAGC-----CCCCCAGCTCTCTTAAGCCCGCGAGGCGCTGTGC--TGCCCGCAGGA 1454

DB 1143 CTTGCTCCAGCTCTTACGACCACTGCGAGAGGGCTGACACCTTTGGCCACACTGAGAA 1202

QY 1455 GCTGGAGAGCGCGAGGAGCTTGGCAAGCGACACTGAAGCGTGCATCTCGGGGAGGA 1514

DB 1203 TGATCAGAGCGCGCAGAGCTGGGGAAGTCCCACTTGGCGAGAGCATGGCTCATCA 1262

QY 1515 GAGGCGCTGCGGCTGACGCGCTGTCGGGTCTGCGGAGGTGACCTGCTTGGGCG 1574

DB 1263 GAGTCTGCTGGAGGCGAGCGCTCTGCAAGAGCTCGGAGCATGAGGTGCTACAGAA 1322

QY 1575 TGTGATCTCCAGACGCTGACGCGCTCACTGCGCAAGTATGCGAGCTCGACCGTGAGGA 1634

DB 1323 GGTGCTGTGTCGCGCACAGGAACGCTCCCGGTCTCAGCTACGCCAGACACAGGAGGAGGA 1382

QY 1635 TGACTTCTGTGAGGCTGCGAGGCGCGGACATCAGCGCTTAAGACCCACAGAGCCAGGA 1694

DB 1383 TGACCTGCTGAACCTGATCGACGCGCCCGCAGCTGTTGTTGCTAAAACTGAGCAAGAGTTGA 1442

QY 1695 GCGCAGGATGCGAGCGCTGCCAGGGGAAAGGGGCTGACATCTTCCATCGGCTGGGCGC 1754

DB 1443 CATTAATTGCCCGCTTCTCAAGCTGACAGTCAAGGACTTCTTCCAGAGAGCTGGGCGC 1502

QY 1755 CTTGCTGTGTCTCAGCCAAAGAACCGGTGGCGGTGGTGGGCGCGCTCCACCTGACCCG 1814

DB 1503 CTTATCTGTGTTTTCGGCTAAACAGCGGTGACGCGCTCTCGAAGCATCCCGTTCACCTGC 1562

QY 1815 AGGAGAGGCGGCTTGGCTCAGCTTTCGGGGAGAGCTGCGCTGTCTCTCATCGCTGCGGT 1874

DB 1563 AGAAGAAAGGAGCTTGGGGTTTCACTTTGAGAGGAAAGCGCCCGCTTTCAGGTTTCACTTCT 1622

QY 1875 CATTCAGGAGCCAGGCGCGGCTGCGCTGAAAGAGGCGGCTACATTTGTGTCACT 1934

DB 1623 GGATCTTATCTGCTTCTGCTCGGTGGCAGGAGCGCGGAAAGAGATTATATTGTCTCCAT 1682

QY 1935 GAATGGGCGAGCATGCGAGGTGGTGGAGCAACGCGAGGTGGTGGAGGAGTGAAGGCTGC 1994

DB 1683 TCAGCTGTGTGATTGTAAGTGGCTGACGCTGAGTGGGTTATGAAGCTGCTGTAAGAGCTT 1742

QY 1995 GGGAGAGCGCGGCGCGAGCTGCGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2034

DB 1743 TGGCGAGGAGGAGATCGAGATGAAGTGAAGTGTGAGGCTGCTGCTGCTGCTGCTGCTG 1782

RESULT 8

AAS15827

ID AAS15827

standard; cdna; 3526 BP.

XX

AC AAS15827;
XX
DT 16-JAN-2002 (first entry)
XX
DE Human ORF for prostate specific protein PSL22.
XX
KW Human; prostate specific protein; PSL22; prostate cancer;
KW benign prostatic hyperplasia; acute prostatitis; testicular cancer;
KW cryptorchidism; testicular disorder; proliferative disorder; lymphoma;
KW leukemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;
KW liver cancer; lung cancer; cytostatic; ss.
XX
OS Homo sapiens.
XX
XX
XX
FH Key Location/Qualifiers
FT CDS l. .2061
FT /*tag= a
FT /product= "PSL22"
XX
PN WO200172962-A2.
XX
XX
XX 04-OCT-2001.
XX
XX 23-MAR-2001; 2001WO-US009410.
XX
XX 24-MAR-2000; 2000US-0191929P.
XX
XX (SAAT/) SAATCIOGLU F.
XX
XX Saatcioglu F;
XX
XX WPI; 2001-662926/76.
XX P-PSDB; AAU10192.
XX
XX New polynucleotide for the diagnosis, prevention and treatment for
XX prostate and testis disorders, particularly prostate cancer, comprises
XX prostate-specific or testis-specific nucleic acids.
XX
XX Claim 5; Fig 14A; 114pp; English.
XX
XX The invention relates to substantially pure prostate-specific or testis-
XX specific polypeptides and the nucleic acids encoding them. Also included
XX are vectors and host cells expressing the proteins, a transgenic animal
XX expressing the protein, antibodies against the proteins, probes for
XX detecting the nucleic acids, antisense molecules for the nucleic acids
XX and methods of isolating modulators of the proteins. Compounds that
XX modulate the prostate specific or testis specific polypeptide are useful
XX to diagnose, prevent or treat disorders of the testis or prostate
XX particularly prostate cancer, benign prostatic hyperplasia, acute
XX prostatitis, testicular cancer, cryptorchidism, undescended, retractile,
XX ascending or vanished testis. Other proliferative disorders for which the
XX modulators may be used include lymphoma, leukaemia, melanoma, ovarian
XX cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer.
XX The present sequence encodes a prostate specific protein, PSL22
XX
SQ Sequence 3526 BP; 992 A; 803 C; 836 G; 895 T; 0 U; 0 Other;

Query Match 20.1%; Score 496.4; DB 5; Length 3526;
Best Local Similarity 56.8%; Pred. No. 2.5e-80;
Matches 977; Conservative 0; Mismatches 731; Indels 12; Gaps 3;

QY 321 GGGCTGTGACTCCCTGACGAGATCCAGTGGCGGACGCTGAGAGCGCGAGGCCAGAT 380
DB 69 GGGCTGTAAATCCCTTGCACAAACCGCGCGGAGTAATTCAGNAATCAAGAGCTGCTT 128
QY 381 TCACGAGAGATTGACAGGAGCTGAGATGCGGACCGGCGCTGAGAACCTTACAGAGC 440
DB 129 GAATCAGCAGATCTCTGAAGCGGTGCGGATGAGGACCGGAGCGGAAACCTTCTGAAAGT 188
QY 441 CACGAGCAACACCGGGTGAGAGAGCGGTGCGGCTGAGCTGAGCTACGTCACACTCAA 500
DB 189 GGCCACAAATCAAAAGGTGGGGAGCAAGTGGCGCTGAGCTGAGCTTGTCAACTCAGA 248

QY 501 CCTGCAGCTGCTGAAGGAGGAGCTGGAGGAGCTCAGCGTGGCGTGGAGACCTGGCCGCGCA 560
DB 249 CCTGCAGATGCTCAAGGAGAGAGCTGGAGGGGCTGAACATCTCGGTGGCGCTTATCAGAA 308
QY 561 TGGGAGCGAAGCTGTCACTGTCCCATGATCCCTCGGCTGAAGGAGACCAAGAGACT 620
DB 309 CACAGAGGAGGCAATTTACGATTCCTTCCTGATTCCTTGGCTGAAGGAAACGAAAGACGT 368
QY 621 GGAATGGTCTACACCGCTGAAGGAGCTGATCTCAGTGCACCTTTTGGAGAGGACGCGCTC 680
DB 369 CGACTTTGGAGTGGCTCTCAAGGATTTTATCTGGAAACATTACAGTGAAGATGGCTATTT 428
QY 681 CTACGAGGCGAGAAATCAGGAGCTGGAGGCTTCGCGCAGGCCCATCGGACCCCGACCG 740
DB 429 ATATGAAGATGAATTTGCAGATCTTATGGATCTGAGACAAAGCTTGTGCGAGCGCTAGCCG 488
QY 741 GAATGAGTGGGCTGGAGCTGTCTACAGCTTATTACAAACAGCTGTGCTTCTTCTGATGC 800
DB 489 GGATGAGGCGCGGGTGGAACTGCTGATGACATCTTCACTCAGCTGGGCTTTGTGCGAGAG 548
QY 801 GCGCTTCTCACCCCTGCCAGAGGCTTCGGGCTCTTCTCCACTGGTACACTGCGTTAC 860
DB 549 TCGATTTCTTCCCGCCACACGCGAGATGGGACTCTGTTCACTGTATGACTCTCTCAC 608
QY 861 TGGGGTCCCGGCCCCAGCAGCGTGCCTTGGGCTTCGAGAGGGGAGCGTTCCTTCAACAT 920
DB 609 TGGGGTTCGGGTGAGCCAGCAGCAACCTGTGCTGGAGAGGCGCAGTGTCTGTTCAACAC 668
QY 921 CGGTGCGCTCCACACGCGATTTGGGGGCGGCCAGGACCGCTCTGACCGAGGGTGGCCG 980
DB 669 TGGGGCGCTCTACACCCAGATTTGGGACCCCGGTGCGATCGGCGACGCGAGCTGGGCTGGA 728
QY 981 CCGCGCTATGAGGCGCTTCCAGAGGCGCTGCGGCTTCCAGCTCTCTGAGGAGAGACTT 1040
DB 729 GAGTGCATAGATGCTTTCAGAGAGCGCGAGGGTTTAAATTAACCTGAAGACACATT 788
QY 1041 CTCCTATGCGCCGAGCCCCAGACATGAGCGCTGCGTCCCTCTGCGCACTGAGAGAGCTCAT 1100
DB 789 TACCCATCTCCAAAGTTACGACATGAGCCCTGCACTGCTCAGCGTGTCTGCTCAAAATGAT 848
QY 1101 GATGGCCAGGCCCCAGGAATGTGTGTTGAGGGCTCTCACACCTTCCCTCCATGGCCCC 1160
DB 849 GCTTGCACAAAGCCCAAGAAAGCGTGTGAGAAATCAGCCTTCCTG-----GATCCG 902
QY 1161 CCAAGACTGCTGGCCAGCTGCGCTGCGCAGGAGGCGCCGAGGTGGCAGCGCAGTA 1220
DB 903 GATGATTTCTTATGCTGTGAGGTGGCTCAGAGGCTGCTAAGGTGGAGAGGTCTA 962
QY 1221 CAGGCTAGTGCAACGAGCCATGCGCCAGCCAGCCCGCTCCAGCTACAGCTAGCTGCTCTCTG 1280
DB 963 CCAACAGCTACACGCGAGCCATGAGCCAGGCGCGGTGAAAGAGAAACATCCCTACTCTG 1022
QY 1281 GACTGCCCTGTGCTGATGTCAGGCGGAGTACTTCCGCTCCCTGGCCCTACACAGTAC 1340
DB 1023 GGCAGCTTAGCTGCGTGAAGGCCCACTACGCGGCGCTGGCCCACTTCTACTGC 1082
QY 1341 CATGGCCCTCTGCGACGGCTCCCGCAGCGACCGGAGAGAGCTCCCGCAGCAGCAGAGGT 1400
DB 1083 CATCTCTCTCATGACACACAGGTGAAGCAGGACCGATCTGGACCAACAGGAGAGTG 1142
QY 1401 CTTCTCTGACG-----CCCCCACTCTCTTAAGCCCGGAGGCCCTGTGTC--TGGCGGAGGA 1454
DB 1143 CTTGTCCAGCTCTACGACCACTGCGAGAGGGGCTGACACCTTGGCCACACTGAAGAA 1202
QY 1455 GCTGGAGGCGGAGCAGCTTGGCAGGACACCTGAAGCGTGCCATCTCTGGGCGAGGA 1514
DB 1203 TGATCAGCAGCGCGCAGCTGGGAGAGTCCCACTTGGCAGAGCCATGGCTCATCAGA 1262
QY 1515 GGAAGGCGCTCGGGTGCACGCGCTGTGCGCGCTCTCTGCGGAGGTGACCTGTCTCGGGC 1574
DB 1263 GGAGTCGGTGGGAGGCGAGCGCTCTGCAAGAGAGCTGCGGAGCATTGAGGTGCTTACAGAA 1322
QY 1575 TGTGATCTCCAGAGCGCTGACGCGCTCACTGGGCCAAGTATGCGGAGCTCGACCGTGGAGGA 1634

CC	ABX97593.1	ABX97008-ABX97185	encode the NOVX proteins described in
CC	ABU65041-ABU65218		
XX			
SQ	Sequence 2109 BP; 518 A; 574 C; 585 G; 432 T; 0 U; 0 Other;		
	Query Match	20.0%;	Score 494.8; DB 6; Length 2109;
	Best Local Similarity	56.7%;	Pred. No. 4.6e-80;
	Matches 976; Conservative	0;	Mismatches 732; Indels 12; Gaps 3
QY	321	GGGCTGTCACTCCCTGACCGCAGATCCAGTGGCGGCGAGCTGCAGAGCCGCGAGGGCCAGAT	380
Db	79	GGGCTGTAATCCCTTGCACAAACGGCGCGAGTAAATTGGAGATCAAGAGCTGCTTT	138
QY	381	TCACAGCAGATTGACAAAGGAGCTGCAGATCGGACGGCGCTGAGAACCTCTACAGAGC	440
Db	139	GAATCAGCAGATCTCTGAAGCCGTGCGGATGAGGACCGGAGCGGAAAAACCTTCTGAAAGT	198
QY	441	CACCAGCAACAAACCGGCTGACAGAGACGGTCCCTGAGAGCTGAGTACGTCAACTCCAA	500
Db	199	GGCCACAAACTCAAAGGTGCGGAGCAAGTCCGGCTGGAGCTGAGCTTCGTCAACTCAGA	258
QY	501	CTTCAGCTGCTGAAGGAGGAGCTGGAGGAGCTCAGCGGTGGCGTGGACCCCTGGGCGGCA	560
Db	259	CTTCAGATGCTCAAGGAGAGCTGGAGGGCTGAAATCTCGGTGGCGTCTATCAGAA	318
QY	561	TGGAGGCGAAGCTGTCTCTGTCCCATGATCCCCCTGGGCTGAAGGAGACCAAGGAGCT	620
Db	319	CACAGAGGAGGCAATTACGATTCCTCTGATTCCTCTTGGCCTGAAGGAAAACGAAAGAGCT	378
QY	621	GGACTGGTTACACCGCTGGAAGGAGCTGATCTCTAGTGCACATTTTCGAGAGGACGGGCGTC	680
Db	379	CGACTTTGCAGTGTCTCTAAGGATTTTATCCTGGAACATTAAGTGAAGATGGCTATTT	438
QY	681	CTACGAGCGAATAACAGGAGCTGGAGGGCCCTCGCGCAGGCCATCGGGACCCCGAGCG	740
Db	439	ATATGAAGATCAAAATTCAGATCTTATGGATCTGAGACAAGCTTGTTCGAGCGCTAGCCG	498
QY	741	GAATGAGTCCGGCTTCGAGCTGCTCAGCGCTATTACAAACAGCTGTGCTTCCTGGAATGC	800
Db	499	GGATGAGGCCCGGGGTGGAATCTGCTGATGACATCTTCATCAGCTGGGCTTTTGTTCGAGAG	558
QY	801	GGCGTTCTCTCACCCCTGCCAGGAGCCTTCGGGCTCTTCTTCACCTGTTACGACTCGCTTAC	860
Db	559	TCGATTTCTTCCGCGCCACACGGCAGATGGGACTCCTGTTACCTGGTATGACTCTCTCAC	618
QY	861	TGGGGTCCCGGCCACGAGCGTGCCTTCGGCTTCGAAAGGCGAGCGTTCTCTTCAACAT	920
Db	619	CGGGGTTCGGTTCAGCCAGCAGAACCTGCTGCTGGAGAAGGCCAGTCTCCTGTTCAACAC	678
QY	921	CGGTGCCCTCCACACGAGATTGGGCGCGCAGGACCGTCTCTGCACGAGGGTGCCTCG	980
Db	679	TGGGGCCCTCTACACCAAGATTGGGACCCCGTGGCATTCGGCAGACGAGCGTGGGCTTGA	738
QY	981	CCGCGCTATGGAGGCGCTTCCAGAGGGCGGCTGGGGCTTCAGCGCTCTGAGGGAGAACTT	1040
Db	739	GAGTGCACATAGATGCTTTTCAGAGAGCGCAGGGGTTTAAATTACCTGAAGACACATT	798
QY	1041	CTCCATGCGCCGAGGCCACGACATGAGCGTGGTCCCTCTGCGCATCTGAGCAGCTCAT	1100
Db	799	TACCCATCTCCAAAGTTTACGACATGAGCCCTGGCATGTCTCAGCGTGTCTGTCAAAATGAT	858
QY	1101	GATGGCCCGAGGCCAGGATGTGTTTGAAGGCGCTCTCACACCTTCCTCATGGCCCC	1160
Db	859	GCTTGCACAAAGCCCAAGAACGGTGTGAGAAAAATCAGCTTTCCTG-----GGATCCG	912
QY	1161	CCAAGACTGCCTGGCCCCAGCTGCGCTGGCGCAGGAGGGCCGCCAGGTGGCAGCCGAGTA	1220
Db	913	GAATGAATTTCTCATGCTGTTGAAGTGGCTCAGGAGGCTGCTAAAGTGGGAGGCTCTA	972
QY	1221	CAGGCTAGTGCCACCGAACCATGGCCACGACCCGCTCCAGCTACGTGCTCTCTGCTG	1280
Db	973	CCAAACAGCTACACGACGACCATGACGAGCGCGCGGTGAAGAGAGAAATCCCTTACTCTCTG	1032

QY 1575 TGTGATCTCCAGACCTGACGGCTCACTGGCCCAAGTATGCGAGCTGACCGTGAGGA 1634
Db 1323 GGTGCTGTGTGCGGCACAGAAAGCTCCCGGCTCACGTACGCTACGCGCAGCACCGAGGAGGA 1382
QY 1635 TGACTTCTGTGAGGCTGCGAGGCCCGGACATCCAGAGCTTAAGACCCACAGAAAGCCAGA 1694
Db 1383 TGACCTGTGAACCTGATGACGCCGCCAGTGTGTGTTGTTAAACTGAGCAAGAGTTGA 1442
QY 1695 GGCACAGATGCCACGGCTGTCCAGAGGAGGGCCCTGACATCTTCATCGGCTGGGGCC 1754
Db 1443 CATATATATGCCCCAGTTCTCCAAAGCTGACAGTCAAGGACTTCTCCAGAAAGCTGGGCC 1502
QY 1755 CTTGCTGTGTTCTACCCAGAACCGGTGGCGCTGGTGGGCCCTCCACCTGACCCG 1814
Db 1503 CTTATCTGTGTTTTCGGCTTAACAGCGGTGGAGCGCTCTCCGAAGCATCGCTTCACTGC 1562
QY 1815 AGAGAGGGCGGCTTTGGGCTCACGCTTCGGGAGACTCGCCTGTCTCATCGCTGCGGT 1874
Db 1563 AGAAGAGGGACTTGGGGTTCACTTTGAGAGGAACGCCCGTTTCAGGTTCACTTCCT 1622
QY 1875 CATTCAGGAGCCAGCGCGCGCGGTGCTGTAAGAGGAGGCGACTACATTTGTGTGAGT 1934
Db 1623 GGATCCTTACTGCTCTGCTCGCTCGGTGGCAGAGCCCGGAAGGAGATTATATTCTCCAT 1682
QY 1935 GAATGGGCGCCATGACGATGTTGGAGACACGCGGAGGTGTTGACGAGCTGAAGGCTGC 1994
Db 1683 TCAGCTTGTGATTGTAAGTGGCTGACGCTGAGTGAAGGTATGAAGCTGCTGAAGAGCTT 1742
QY 1995 GGAGAGGGGGCGCGCCAGGCTTCAGGTGTTGTGCTGCTGCTG 2034
Db 1743 TGGCGAGGACGAGATCGAGATGAAGTCGTGAGCCTCTG 1782

RESULT 11

ABSF71855
ID ABSF71855 standard; cDNA; 3484 BP.
XX ABSF71855;
AC ABSF71855;
XX
DT 02-DEC-2002 (first entry)
XX
DE Human cDNA encoding GTP-Rho binding protein 2.
XX
KW Human; ss; GTP-Rho binding protein 2; GRBP2; chromosome 19q12; oncogene;
KW tumour; liposarcoma; ichthyosis congenita III;
KW benign familial infantile convulsion; gene therapy.
XX

OS Homo sapiens.
XX
XX EP1231216-A2.
XX
XX
PD 14-AUG-2002.
XX
PF 17-JAN-2002; 2002EP-00001026.
XX
XX 30-JAN-2001; 2001WO-US000663.
PR 30-JAN-2001; 2001WO-US000664.
PR 30-JAN-2001; 2001WO-US000665.
PR 30-JAN-2001; 2001WO-US000666.
PR 30-JAN-2001; 2001WO-US000667.
PR 30-JAN-2001; 2001WO-US000668.
PR 30-JAN-2001; 2001WO-US000669.
PR 30-JAN-2001; 2001WO-US000670.
PR 29-JUN-2001; 2001US-00895040.
XX (AEOM-) AEOMICA INC.
XX
XX Shannon ME, J1 Y;
PI
XX WPI; 2002-684026/74.
DR P-PSDB; ABG94709.
XX
XX

PT Novel GTP-Rho binding protein 2 and nucleic acids encoding the protein,
PT useful for the manufacture of a medicament for treating a disease
XX associated with altered expression or activity of human GRBP2 protein.
PS Claim 1; Fig 3; 101pp; English.

XX The invention relates to an isolated GTP-Rho binding protein 2 (GRBP2)
CC polypeptide or a fragment of at least 6 amino acids or a sequence in
CC which at least 95% of deviations from GRBP2 sequences are conservative
CC substitutions. Also included are an isolated nucleic acid (GRBP2 NA)
CC encoding GRBP2 comprising the full length cDNA or CDS, fragments or
CC variants, GRBP2 vectors, host cells, antibodies, transgenic non-human
CC animals modified to contain GRBP2 NA for unable to express the endogenous
CC orthologue of GRBP2), diagnosing a disease caused by a mutation in human
CC GRBP2 or altered expression of GRBP2, ant-agonists of GRBP2, GRBP2
CC microarrays, fusion proteins and screening for agents that modulate the
CC expression of GRBP2 NA. GRBP2 is useful for identifying binding partners
CC of GRBP2. GRBP2, GRBP2 NA and Ab are useful in therapy and in the
CC manufacture of a medicament for the treatment or prevention of a disorder
CC associated with increased or decreased expression or activity of human
CC GRBP2 (e.g. tumours, liposarcoma, ichthyosis congenita III and benign
CC familial infantile convulsion, all associated with the chromosomal
CC location of GRBP2, 19q12). GRBP2 is useful as a standard in immunoassay
CC specific for the proteins, to be used in a therapeutic agent, as
CC vaccines, to be and as antigens (e.g. for epitope mapping) or immunogens
CC (e.g. for raising antibodies. GRBP2 NA is useful as hybridisation probes,
CC to prime synthesis of nucleic acids, to prime first strand cDNA sequence,
CC on an mRNA template, and to drive in vivo expression of the proteins. The
CC vector is useful for shuttling GRBP2 NA between host cells derived from
CC disparate organisms, for inserting GRBP2 NA into host cell chromosome,
CC for expressing sense or antisense RNA transcripts of GRBP2 NA in vitro or
CC within a host cell, and for expressing GRBP2 alone or as fusions to
CC heterologous polypeptides. The antibody is useful as an analytical
CC reagent for detection and quantification of GRBP2 and as an immuno
CC therapeutic agent and is useful for flow cytometric detection, for
CC scanning laser cytometric detection, or for fluorescent immunoassay. The
CC present sequence is a GRBP2 cDNA sequence
XX
XX Sequence 3484 BP; 932 A; 814 C; 842 G; 896 T; 0 U; 0 Other;

Query Match 20.0%; Score 493.2; DB 6; Length 3484;
Best Local Similarity 56.7%; Pred. No. 9.4e-80;
Matches 975; Conservative 0; Mismatches 733; Indels 12; Gaps 3;

QY 321 GGGCTGTGACTCCTGACGAGATCCAGTGGCGCCAGCTGCAGAGCGCGAGGCCAGAT 380
Db 89 GGGCTGTAAATCCCTTGCACAAACCGCGGAGTAATTCGAGAAACAAGAGCTGCTTT 148
QY 381 TCACACGAGATGTACAAGGAGCTGCAGATGCGAGCGGGCGCTGAGAACCTCTACAGAGC 440
Db 149 GAATCAGCAGATCCTGAAGCCGTGCGGATGAGATCGGAGCGGAAACCTCTGAAAGT 208
QY 441 CACGACAAACACCGGCTGAGAGAGACGGTGGCCCTGGAGCTGAGCTACGTCAACTCCAA 500
Db 209 GGCACAAAATCAAAAGGTGCGGAGCAAGTGGCGGCTGGAGCTGAGCTTCTCAACTCAG 268
QY 501 CCTGCAGCTGTGAAGGAGGAGCTGGAGGAGCTCAGCGGTGGCGTGACCTCGCCGCGCA 560
Db 269 CCTGCAGATGCTCAAGGAAGAGCTGGAGGGCTGGAACATCTCGGTGGCGGTCTATCAGAA 328
QY 561 TGGGAGCGAAGCTGTCACTGTCCCATGATGCCCTGGGCGCTGAAGAGACCAAGAGCT 620
Db 329 CACAGAGGAGGCAATTTACGATTCCCTGTCTCTTGGCCTGAAGAAACGAAAGACGT 388
QY 621 GGACTGCTTACACCGCTCAAGGAGCTGATCTCAGTGCCTTTGGAGAGACGCGGCCCTC 680
Db 389 CGACTTTGCAAGTGTCTCTCAAGGATTTTATCTCGAACATTACAGTGAAGATGGCTATT 448
QY 681 CTACGAGGCGAGAAATCAGGAGGCTGGAGGCCCTTGGCGGAGGCCATCGGACCCCGAGCCG 740
Db 449 ATATGAAGATGAATAATGTCAGATCTTATGATCTGAGACAGAGCTTTGTGCGAGCGCTAGCCG 508
QY 741 GAATGAGTCGGGCTGGAGCTGCTCAGCCTATTACACCTATTACACAGCTGTGCTTCTGATGC 800

509	Db	GGATGAGGCGGGGTGGAATGCTGATGACATATTATCAGCTGGGCTTTGTTCGAGAG	568
801	Qy	CGCTTCTCTACCCCTCGCAGAGAGCTCGGGCTCTTCTTCACCTGGTACGACTCGCTTAC	860
569	Db	TCGATTCTTCGGCCACACGCGAGATGGGACTCTGTTCACCTGGTATGACTCTCTCAC	628
861	Qy	TGGGGTCCGGCCCGACGAGCGTGCTTCGGCCCTTCGAGAAGGCGAGCGTCTCTTTCAACAT	920
629	Db	CGGGTTCGGTTCAGCCAGCAGAACTCTCTGCTGGAAGGCCAGTGTCTGTTCACAC	688
921	Qy	CGGTCCCTCCACACGCGAGATTGGGCGCGCCAGGACGGCTCTCTGCAACGAGGTGCCCG	980
689	Db	TGGGGCCCTCTACACCCAGATTGGGACCCGGTGGCATCGGCGACGCGAGGCTGGGGTGA	748
981	Qy	CCGCGCTATGAGGCGCTTCCAGAGGGCGCTGGGGCCCTTCAGCCCTCTGAGGGAGAACTT	1040
749	Db	GAGTGCATAGATGCCITTCAGAGAGCGCAGGGGTTTAAATTACCTGAAGACACATT	808
1041	Qy	CTCCATGCGCCGAGCCGACAGATGAGGGTGTGCTCTCTGTGGCACTGAGAGAGTTCAT	1100
809	Db	TACCCATACTCCAAGTTACGACTGAGCGCTGCCATGCTCAGCGTGTCTGCTCAAAATGAT	868
1101	Qy	GATGGCCAGGCCACAGAAATGTGTGTTGAGGGCCCTTCACCACTTCCTCCATGGCCCC	1160
869	Db	GCTTGCACAGCCCAAGAAACGGTGTGAGAAATCAGCGTTCCTG-----GGATCCG	922
1161	Qy	CAAAGACTGCCCTGGCCCGCAGCTGCGCCTGGGCGAGGGCGCCGAGCTGGCGAGCCAGTA	1220
923	Db	GAATGAATCTTATCTGTGTGAAGTGGCTCAGAGGCTGCTAAGGTGGGAGGTCCTA	982
1221	Qy	CAGGCTAGTGCAACGGAACATGCGGCCAGCCACCCTCCACGACTAGTGGCTGTCTCTG	1280
983	Db	CCAAAGACTACACGAGCCATGAGCCAGGGCGCGGTGAAGAGAAATCCCTTACTCTCG	1042
1281	Qy	GACTGCCCTGTGTGATGTCAGAGCCGAGTACTTCGGCTCCCTGGCCCACTTACCACTGAGC	1340
1043	Db	GGCCAGCTTTAGCCTGCGTGAAGGCCCAACCACTAGCGGCGCTGGCCCACTTCTACTGC	1102
1341	Qy	CATGGCCCTCTGCGACGGCTCCCCAGCGACCGAGGGAGAGCTCCCAAGCACGAGCAGGT	1400
1103	Db	CATCTCTCATCGACCAACAGGTGAAGCCAGGACCGATCTGGACACACGAGAGAATG	1162
1401	Qy	CTTCTCTGAGC-----CCCCACCTCTCTTAAGCCCCGAGGCCCTGTGC--TGCCGCGAGGA	1454
1163	Db	CTGTCCAGCTCTPACGACCACATGCGAGGGGCTGACACCTTGGCCACACTGAAGAA	1222
1455	Qy	GCTGGAGGCGGACGAGCTTGGCANGGCACACTGAAGGGTGCCATCTGGGGCGGGA	1514
1223	Db	TGATCAGACGCGCGACAGCTGGGGAAAGTCCCCACTTCGCGAGAGCCATGCTCATACGA	1282
1515	Qy	GGAGGCGCTCGGCTGACGCGCTGTGCCGCTCTGCGGAGGTGACCTGCTTCGGGC	1574
1283	Db	GGAGTCGTGGGAGGCAAGCTCTGCAAGAGCTGCGAGGATTTAGGTGTCTACAGAA	1342
1575	Qy	TGTGATCTCCAGACGCTGCAGCGCTCACTGGCCAGTATGCCGAGCTCGACCGTGAGGA	1634
1343	Db	GGTCTGTGTGCCGACAGGAAACGCTCCCGCTCACGTAACCCAGCACGAGGAGGA	1402
1635	Qy	TGACTTCTGTGAGGCTCGGAGGCCCGGACATCCAGCTTAAGNCCACAGAAAGCCAGA	1694
1403	Db	TGACCTGCTGAACCTGATCGAAGCGCCCGACGTGTTGTTGCTTAAACTGAGCAAGAGGTGA	1462
1695	Qy	GGCCAGGATGCCAGCTGTCTCCAGGGGAAGGGCCCTGCATCTTCATCGGCTGGGGCC	1754
1463	Db	CATTATATTGCCCGAGTTCTCCAGCTGCAGTACGGAATTTCTTCGAAGCTGGGCC	1522
1755	Qy	CCTCTCTGTGTTCTCAGCCAAGAACCGGTGGCGGCTGGTGGGGCCCGTCCACCTGACCCG	1814
1523	Db	CTTATCTGTGTTTTTCGGCTAACAGCGGTGGACGCGCTCCTCGAAGCATCGCTTCACTGC	1582
1815	Qy	AGGAGGCGGCGCTTTGGCTCAGCTTTCGGGGAGACTCGGCTGTCTCATTCGCTGCGGT	1874

Db	1593	AGAGAGGGGACTGTGGGTTTCACCTTTGAGAGGGAAACCCCGGTTTCAGGTTTCACCTTCCT	1644
Qy	1875	CATTCCAGGAGCAGCCAGCGCGCGCTGGCCTGAAGGAGGGGACTACATTGTGTCACT	1934
Db	1643	GGATCCCTTACTGCTCTCGCTCGGTGGGAGGAGCCCGGAGGAGATTATATTGTCTCAT	1702
Qy	1935	GAATGGGGAGCCATCGAGGTGGTGAGACACGCGGAGGTGGTGACGAGCTGAAGGCTGC	1994
Db	1703	TCAGCTGTGTGATTGTAAGTGGCTGACGCTGAGTGAGGTTATGAAGCTGCTGAAGAGCTT	1762
Qy	1995	GGGAGAGCGCGCGCCAGCCCTGCAGGTGGTGTGCTGCTCTG	2034
Db	1763	TGGGAGGACGAGATCGAGATGAAGTCGTGAGCCTCTCTG	1802
RESULT 12			
AAFS58361			
ID	AAFS58361 standard; cDNA, 3019 BP.		
XX	AC	AAFS58361;	
XX	DT	19-APR-2001 (first entry)	
XX	DE	Human GTP-binding associated protein #61 coding sequence.	
KW	KW	Human; guanosine triphosphate binding associated protein; GTP; GBAP; inflammation; AIDS; Addison's disease; anaemia; arteriosclerosis; asthma; autoimmune disorder; hepatitis; multiple sclerosis; cancer; diabetes; osteoporosis; psoriasis; ss.	
OS	OS	Homo sapiens.	
XX	XX	WO200105970-A2.	
PN	PN	25-JAN-2001.	
PD	PD	19-JUL-2000; 2000WO-US019698.	
PF	PF	19-JUL-1999; 99US-0144595P.	
XX	XX	23-AUG-1999; 99US-0150460P.	
PR	PR	15-OCT-1999; 99US-0159849P.	
XX	XX	(INCY-) INCYTE GENOMICS INC.	
PA	PA	Yue H, Tang YT, Bandman O, Hillman JL, Lal P, Au-Young J; Reddy R, Yang J, Baughn MR, Lu DM, Azimzai Y, Patterson C; WPI; 2001-091972/10.	
DR	DR	P-PSDB; AAB68561.	
XX	XX	New guanosine triphosphate-binding associated proteins (GBAP) and their encoding nucleic acids, useful for treating and/or diagnosing diseases associated with GBAP expression, such as cancer, diabetes and asthma.	
PT	PT	Claim 5; Page 229-230; 233pp; English.	
XX	XX	The present invention relates to novel human guanosine triphosphate (GTP) -binding associated proteins (GBAPs; AAB68501-AAB68566) and their coding sequences (AAFS58301-AAFS58366). The proteins and coding sequences of the present invention are useful for treating a variety of disorders including inflammation, AIDS, Addison's disease, anaemia, arteriosclerosis, asthma, autoimmune disorders, Grave's disease, hepatitis, multiple sclerosis, cancer, diabetes, osteoporosis and psoriasis	
XX	SQ	Sequence 3019 BP; 800 A; 708 C; 762 G; 748 T; 0 U; 1 Other;	
Query Match			
Best Local Similarity 56.9%; Score 487.8; DB 4; Length 3019;			
Matches 957; Conservative 0; Mismatches 713; Indels 12; Gaps 3;			
Qy	359	TGCAGAGCCGAGGGCCAGATTCCACGAGATTGCAAGGAGCTGCAGATGCGGACGG	418

```
Db 190 TGCAAGATCAAGAGAGCTGCTTTTGAATCAGCAGATCCTGAAAGCCGTGCGGATCAGGACCG 249
Qy 419 GCGCTGAGAACCTCTACAGAGCCACCAGCAACAAACCGGTGAGAGACGCTCGCCCTGG 478
Db 250 GAGCGGAAAACCTTCTGAAAGTGCGGCACAACTCAAAAGGTGCGGAGCAAGTTCGGCTGG 309
Qy 479 AGCTGAGCTACGTCAACTCAACTCAACTCAGCTGCTGAAGGAGGAGCTGGAGGAGCTCAGCG 538
Db 310 AGCTGAGCTTCGTCAACTCAGCTCAGATGCTCAAGGAGAGCTGGAGGAGCTGAACA 369
Qy 539 GTGGCTGGACCCCTGGCGGCATGGGAGCGAAGCTGTCACTGTCCCATGATCCCTCTGG 598
Db 370 TCTCGGTGGCGCTCTATCAGAAACACAGAGGAGCAATTTACGATTCCTCCCTGATTCCTTG 429
Qy 599 GCTTGAAGGAGAACAGGAGCTGAGCTGCTACACCGCTGAAGGAGGAGCTGATCTCAGTGC 658
Db 430 GCCTGAAGGAAAAGAAAGAGCTGCACTTTTGCACTCGTCTCTCAAGGATTTTATCCTGGAAC 489
Qy 659 ACTTTGGAGAGGACGGCGCTCCTACGAGGACGAAATCAGGAGGAGCTGGAGGCGCTTCGGC 718
Db 490 ATTACAGTGNAGATGGCTATTATATGAAGATGAAATTCAGATCTTATGGATCTGAGAC 549
Qy 719 AGCCCATGCGGACCCCGAGCGGAATGAGTCGGGCTGGAGCTGTCTCAGCCTATTTACA 778
Db 550 AAGCTTGTGCGAGCGCTAGCCGGGATGAGCGCGGGTGGAACTGCTGATGACATACCTCA 609
Qy 779 ACCAGCTGTGCTTCTGGATGCGCGCTTCTCACCCTGCGAGGAGCTCGGCTCTTCT 838
Db 610 TCCAGCTGGGCTTTGTGAGAGTGCATTTCTTCGCGCCACACCGGCAAGTGGGACTCTGT 669
Qy 839 TCCACTGGTACGACTCGCTTACTGGGCTCCCGGCCAGCAGCGCTGCCCTGGCTTCGAGA 898
Db 670 TCACCTGGTATGACTCTCTACCGGGGTTCCGGTCAGCAGCAGAACCTGCTCTGGAGA 729
Qy 899 AGGCGACGCTTCTTCAAATCGGTGCCCTCCACAGCGAGATGGGGCGCGCAGAAC 958
Db 730 AGGCCAGTGTCTGTTCAACACTGGGGCCCTCTACACCCAGATGGGACCGGTGTGATC 789
Qy 959 GCTCTCTGCACCGAGGGTGGCGCGGCTATGAGGCTTCCAGAGGCGCGCTGGGCGCT 1018
Db 790 GGCAGACGAGGTGGGCTGGAGTGCCATAGTGCCTTTTCAGAGAGCGCGAGGGGTTT 849
Qy 1019 TCAGCTCTCTGAGGAGAACTTCTCCCATGCGCGAGCCGACAGCATGAGCGCTGCGTCCC 1078
Db 850 TAAATTACCTGAAGACACATTTACCCATCTCCAGTTACGATCAGATGAGCCCTGCCATGC 909
Qy 1079 TCTGGGACCTGGAGCAGCTCATGATGCCCGGCGGAGGAAATGTGTGTTTGAAGGCGCTCT 1138
Db 910 TCAGCGTGTCTGTCAAAATGATGTTTGCAAGGCGCCAAAGAAAGCGTGTGAGAAAATCA 969
Qy 1139 CACCACTGCTCCATGGCCCGCCCAAGACTGCTGGCCAGCTGCGCTCGCGCAGGAGG 1198
Db 970 GCCTTCTG-----GGATCCNGAATGAATCTTTCATGCTGGTGAAGGTGGCTCAGGAGG 1023
Qy 1199 CCGCCAGGTGGGAGCGGAGTACAGGCTAGTGACCGGACCATGGCCGACCCACCCGCTCC 1258
Db 1024 CTGCTAAGTGGGAGAGTCTACCAACAGCTACACGCGACCATGAGCCAGCGCGGTGA 1083
Qy 1259 ACAGCTACGTGCTGTCTCTGAGCTGCCCTGGTGCAATGTCAAGCGCGAGTACTTCCGCT 1318
Db 1084 AAGAGAACTACCCCTACTCTCTGGGCCAGCTTAGCTGCTGTAAGGGCCCAACCTACGCG 1143
Qy 1319 CCTGGGCCACTACACGATGACCATGGCCCTCTGCGACCGCTCCCGACGACGAGGAG 1378
Db 1144 CCTGGGCCACTACTTCACTGGCCATCTCTCTCATCGACCAAGGTTGAAGCCAGGACGG 1203
Qy 1379 AGCTCCCAACGACGAGCAGGTCTTCTCTGAGC-----CCGCCACTCTCTTAAGCCCGCA 1434
Db 1204 ATCTGGACCAACGAGGAAAGTGCCTGTCACAGCTCTACGACCAATGCCAGAGGGGTGA 1263
Qy 1435 GGCCCTGTGC--TGCCCGCAGGAGCTGGAGAGCGCAGGAGCTGGCAAGGCACACCTGA 1492
Db 1264 CACCCCTGGCCACACTGAAGAAATGATCAGCAGCGCGCACAGCTGGGGAATGCCACTTGC 1323
```

```
Qy 1493 AGCGTCCATCTCTGGGCGCAGGAGGAGCGCTGCGGCTGCACGCGCTGTGCGCGTCTCTGC 1552
Db 1324 GCAGACCATTGGCTCATACAGAGAGTGGTTCGGGAGCGGAGCGCTCTGCAAGAAGCTGC 1383
Qy 1553 GCAGGTGAGACCTGCTTCGGGCTGTGATCTCCAGACGCTGCAGCGCTCACTGGCCAAAGT 1612
Db 1384 GGACGATTGAGGTGCTTACAGAGGTGCTGTGTGCGCAGAGAACGCTCCCGGCTCAGCT 1443
Qy 1613 ATGCGGAGCTCGACCGTGTGAGGATGACTTCTGTGAGGCTGCCAGGCCCCCGGACATCCAGC 1672
Db 1444 ACGCCAGCAGCACGAGGAGGAGTACCTGTGAACCTGATCGACGCGCCCCCAGTGTGTG 1503
Qy 1673 CTAAGACCCACCAAGAACGAGGCGCAGGATGCCAGCTGTCCAGGGGAAGGGGCGCTG 1732
Db 1504 CTAAACTGAGCAGAGGTTGACATTAATTATGCCAGTTCCTCAAGCTGACAGTCAACGG 1563
Qy 1733 ACATTTCCATCGGCTGGGCGCCCTCTGTGTGTTCTCAGCAAGAACCGGTGCGCGGTGG 1792
Db 1564 ACTTCTTCCAGAACTGGGCGCCCTTATCTGTGTTTTCGGCTAACAAAGCGGTGACGCGCTC 1623
Qy 1793 TGGGGCCCGTCCACCTGACCCGAGGAGAGCGGCGCTTGGCCTCAGCTTCGGGGAGACT 1852
Db 1624 CTCGAAAGCATTCGCTTCACTGCAAGAAAGAGGAGCTTGGGGTTTCACTTGAGAGGGAACG 1683
Qy 1853 CGCTGTCTCATCGCTGCGCTCATTCAGGGAGCCAGGCGCGCGCGGCTGGGCTGAAGG 1912
Db 1684 CCCCCTGTGAGTTCACCTTCTGATCCCTTACTGTCTGCTGCTGGTGGCAGGAGCCCGG 1743
Qy 1913 AGGCGCACTACATTTGTGTGAGTGAATGGGAGCCATGCGAGGTGGTGGAGACACGCGGAGG 1972
Db 1744 AAGGAGATTATATTGTCTCCATTGCTGATTTGTAAGTGGCTGACGCTGAGTGAGG 1803
Qy 1973 TGGTGAAGGAGCTGAAGGCTGCGGAGAGCGGCGCGGCGGCTGCAAGTGGTGTGCTGTCG 2032
Db 1804 TTATGAAGCTGCTGAAGAGCTTTGGCGAGGACGAGATCGAGATGAAGTCTGTGAGCGCTCC 1863
Qy 2033 TG 2034
Db 1864 TG 1865

RESULT 13
ABX97182
ID ABX97182 standard; cdna; 2310 BP.
AC ABX97182;
XX
DT 20-MAY-2003 (first entry)
XX
DE Human NOV126a cDNA.
XX
KW NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200272757-A2.
XX
PD 19-SEP-2002.
XX
PF 08-MAR-2002; 2002WO-US0006908.
XX
PR 08-MAR-2001; 2001US-0274101P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 08-MAR-2001; 2001US-0274322P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
```


PR 14-MAR-2001; 2001US-0276000P.
ER 16-MAR-2001; 2001US-0276776P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 28-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0277338P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.
XX (CURA-) CURAGEN CORP.
EA Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
XX Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
PI Fernandes ER, Casman ES, Malyankar UM, Gerlach V, Liu Y, Anderson D;
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
PI Lepley DM, Rieger DK;
XX WPI; 2002-723332/78.
DR P-PSDB; ABU65215.
XX
XX NOVX polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVX expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
PT asthma.
XX
PS Claim 13; Page 470; 1103pp; English.

XX This invention describes novel human NOVX polypeptides which have
CC cytosolic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive
CC activity. pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABX97008-ABX97185 are cDNA fragments amplified and isolated by
CC the PCR primers and probes represented in ABX13460-ABX13462 and ABX97186-
CC ABX97593. ABX97008-ABX97185 encode the NOVX proteins described in
CC ABU65041-ABU65218
XX
SQ Sequence 2310 BP; 569 A; 622 C; 628 G; 491 T; 0 U; 0 Other;
Query Match 18.8%; Score 463.6; DB 6; Length 2310;
Best Local Similarity 56.7%; Pred. No. 2e-74;
Matches 977; Conservative 0; Mismatches 729; Indels 16; Gaps 6;
OY 321 GGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGCCAGCTGCAGAGCCGAGGCCAGAT 380
DB 114 GGGCTGTAAATCCCTTGCACAAACCGCGCGAGTAATTTGCAGAAATCAAGAGCTGCTTT 173
OY 381 TCACGAGCAGATTGACAAAGGAGCTGCAGATCGGAGCGGCGCTGAGAACCTTCTCAGAGC 440
DB 174 GAATCAGCAGATCTCTGAAAGCCGTGCGGATGAGGACCGGAGCGGAAACCTTCTGAAAGT 233
OY 441 CACCAGCAACAAACCGGGTGAGAGAGCGTCCCTCGAGCTGAGTACGTCACTCCAA 500
DB 234 GGCCACAACCTCAAGGTGCGGAGCAAGTCCGGCTGGAGCTGAGCTTTCGTAACCTCAGA 293
OY 501 CCTGCAGTCTGTAAGGAGGAGCTGGAGAGCTCAGCGGTGGCGTGGAGCCCTGGCGCGCA 560
DB 294 CCTGCAGATGCTCAAGGAAGAGCTGGAGGGGCTGCAACATCTCGGTGGCGCTTATCAGAA 353
OY 561 TGGGAGCGAAGCTGTCACTGTCCCATGATCCCTCGGCTGAGGAGCAAGAGAGCT 620
DB 354 CACAGAGGAGGCAATTTACGATTCCTCTTGGCTTGAAGGAACGAAAGAGCT 413
OY 621 GGACTGGTCTACACCGCTGAAGGAGCTGATCTCAGTGCACCTTTGGAGAGGAGCGGCTC 680
DB 414 CGACTTTCAGTCTGCTTCAAGGATTTTATCTGGAACATTACAGTGAAGATGGCTATTT 473
OY 681 CTACGAGCAGAAATCAGGAGAGCTGGAGGCGCTCGCGCAGGCGCATCGGAGCCCGCAGCG 740
DB 474 ATATGAAGATGAATTTGCAGATCTTATGGATCTGAGCAAGCTTGTTCGACCGCTAGCG 533
OY 741 GAATGAGTCGGGCTTGGAGCTGCTCAGAGCTTATACAGCCTATTTACAGCAGCTGTCTCTG 800
DB 534 GGATGAGCGCGGGTGGAACTGCTGATGACATATCTTCATCAGCTGGGCTTTGTGAGAG 593
OY 801 GCGCTTCTCACCCTGCGCAGGAGCTCGGGCTCTTCTTCCACTGGTACGACTGCTTAC 860
DB 594 TCGATTCTTCCGCGCCACACGCGAGATGGGACTCTGTTCACCTGGTATGATCTCTCAC 653
OY 861 TGGGCTCCCGCCAGCAGCGTGCCTGGCGCTTCGAGAAGGCGCAGCGCTTCTTCAACAT 920
DB 654 CCGGCTTCCCGCTCAGCCAGCAGAACTGCTGCTGGAGAAGGCCAGTGTCTTGTTCACAC 713
OY 921 CGGTGCGCTTCCACAGCAGATTTGGGCGCGCCAGCAGCGCTTCTGACCGAGGCTGCCG 980
DB 714 TGGGCGCTTCTACACCCAGATTTGGGAGCCCGGTGTGATCGGAGCAGCGAGCTGGGCTG 773
OY 981 CCGCGCTATGAGGCGCTTCCAGAGGCGCTGGGCGCTTTCAGCGCTTCTGAGGAGAACTT 1040
DB 774 GAGTGCCATAGATGCCCTTTCAGAGCGCGCAGGGGTTTAAATTTACCTGAAGAAGACAT 833
OY 1041 CTCCTCATGCGCGAGCCAGCAGATGAGCGCTGCGTCCCTCTGCGCACTGGAGAGCTCAT 1100
DB 834 TACCCATCTCCAAGTTACGACATGAGCCCTGCCATGCTCAGCGTGTCTCGTCAAAATGAT 893
OY 1101 GATGCGCCAGCCCGAGGATGTGTTTGGGGCTCTTCACCACTCGCTCCATCGGCC 1160

Db 894 GCTTGACAAAGCCAAAGAACGCTGTTTGAGAAAATCAGCCTTCCTG-----GGATCCG 947
Qy 1161 CAAAGACTGCTGGCCAGCTGCGCTTGGCGAGAGAGCGCCAGGTGCGACCGAGTA 1220
Db 948 GAATGAATTTTCATGCTGGTGAAGGTGGCTCAGAGAGGCTGCTAAGGTGGAGAGTCTA 1007
Qy 1221 CAGGCTAGTGACCGGACCATGCCCAGCCACCCGTCACGACTACGTGCTGCTTCCTG 1280
Db 1008 CCAACAGCTACACGACGATAGCAGGCGCCGCGTGAAGAGAAATCCCTTACTTCCTG 1067
Qy 1281 GACTGCCCTGGTGATGTCAAGCCGAGTACTTCCGCTCCCTGGCCCACTACACAGTACG 1340
Db 1068 GGCAGCTTAGCTGGTGAAGCCACACACTAGCGGCGCTGGCCCACTTCACTTCG 1127
Qy 1341 CATGGCCCTTGGAAGGCTCCAGGACCGAGGAGAGCTCCCCAGCCACAGCAGGT 1400
Db 1128 CATCTCTCTATCGACCAACAGCTGAAGCCAGGACCGGATCTGGACCAACAGGAGGTG 1187
Qy 1401 CTTCTCTGACG-----CCCCACCTCTCTTAAGCCCGAGGCGCTGTGC--TGCGGAGGA 1454
Db 1188 CCGTCTCCAGCTCTACGACCAATGCGAGGGGCTGACACCCCTTGGCCACACTGAGAA 1247
Qy 1455 GCTGGAGGAGCGCAGGAGCTTGCGCAAGGCACACCTGAAGCGTGCCATCCTGGGCGAGGA 1514
Db 1248 TGATCAGCAGCGCGCAGCTGGGGAAGTCCCACTTGGCAGAGCCATGGCTCATCAGGA 1307
Qy 1515 GAGGCGCTGCGCTGACCGCCCTGTGCCCGCTCTCGCGGAGGTGACCTGCTTCGGGC 1574
Db 1308 GGAAGTGGTGGGAGGCGCAGCTCTGCAAGAGAGCTGGGAGCAATTGAGTGTCTACAGAA 1367
Qy 1575 TGTGATCTCCAGACGCTGCGAGCGCTCACTGGGCCAAAGTATGCGAGCTGCACCGTGAAGGA 1634
Db 1368 GTGCTGTGTGCGCGCAGGAAAGCTCCCGCTCACTAGCGCCAGCACCAGGAGAGGA 1427
Qy 1635 TGACTTCTGAGGCTGCGAGGCGCC--GGAATCAGCGCTAAGACCCACACAGAGCCCA 1692
Db 1428 TGACCTGCTGAACCTGATCGACGCGCCCGCAGAGTGTGTTGCTAAACCTGAGCAAGAGTT 1487
Qy 1693 GAGGCCAGGATGCCAGCCTGTCCAGGGAAGGGCGCTGACATCTTCCATCGCTGGGG 1752
Db 1488 GACATTATATGCGCCAGATTCT--CCAGCTACAGTCAOGGACTTCTTCCAGAGCT--GGG 1545
Qy 1753 CCCTGCTGTGTTCTCAGCAAGAACCGGTGGCGGTGTTGGGGCGCTGCCACCTGAGCC 1812
Db 1546 CCCTTATCTGTGCTGCGCTACAGCGGTGAGCGCTCTCGAAGCATCCGCTTCACT 1605
Qy 1813 CGAGGAGAGGCGGCTTTGGCCTACGCTTTCGGGAGACTCGCTGTGCTCTCATCGCTGCC 1872
Db 1606 GCAGAAAGAGGGACTTGGGGTTACCTTGAGAGGGAACGCCGCCGCTTCAAGTTCACTTC 1665
Qy 1873 GTCAATTCAGGAGCCAGCGCGCGGCTGAGAGGAGGCGGCTGACATTTGCTCA 1932
Db 1666 CTGGATCTTACTGTCTCTGCTCGGTGGAGGAGCCCGGGAAGGAAATATATGCTCTC 1725
Qy 1933 GTGAATGGGAGCCATGCAAGTGGTGAGACACCGCGAGGTGGTGAGCGAGCTGAAGGCT 1992
Db 1726 ATTCACTTGTGATTTGATGCTGACGCTGAGTGAGGTATGAGCTGCTGAGAGC 1785
Qy 1993 GCGGAGAGCGGCGCGCAGCTGCGAGGTGGTGTGCTGCTG 2034
Db 1786 TTTGGGAGGACGAGATCGAGATGAAGTGTGAGGCTCTCTG 1827

RESULT 14

AAC91349 standard; cDNA; 2757 BP.

AC AAC91349;

DT 16-MAR-2001 (first entry)

DE Human polynucleotide for diagnostics and therapeutics, SEQ ID NO: 49.
XX

KW Human; diagnostics and therapeutics; dithp; cytostatic;
KW immunosuppressive; nootropic; neuroprotective; antiviral; antiallergic;
KW hepatotropic; antidiabetic; antiinflammatory; antiulcer; vulnerary;
KW anticonvulsant; antibacterial; antifungal; antiparasitic; cardiant;
KW cancer; immune disorder; cardiovascular disorder; neurological disease;
KW infection; endocrine disorder; metabolic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200073509-A2.
XX
PD 07-DEC-2000.
XX
XX 31-MAY-2000; 2000WO-US015404.
XX
PR 01-JUN-1999; 99US-0137161P.
PR 02-JUN-1999; 99US-0137109P.
PR 02-JUN-1999; 99US-0137113P.
PR 02-JUN-1999; 99US-0137114P.
PR 02-JUN-1999; 99US-0137173P.
PR 02-JUN-1999; 99US-0137258P.
PR 02-JUN-1999; 99US-0137259P.
PR 02-JUN-1999; 99US-0137260P.
PR 03-JUN-1999; 99US-0137337P.
PR 03-JUN-1999; 99US-0137396P.
PR 03-JUN-1999; 99US-0137411P.
PR 03-JUN-1999; 99US-0137417P.
PR 04-AUG-1999; 99US-0147377P.
PR 04-AUG-1999; 99US-0147436P.
PR 05-AUG-1999; 99US-0147500P.
PR 05-AUG-1999; 99US-0147520P.
PR 05-AUG-1999; 99US-0147527P.
PR 05-AUG-1999; 99US-0147530P.
PR 05-AUG-1999; 99US-0147536P.
PR 05-AUG-1999; 99US-0147541P.
PR 05-AUG-1999; 99US-0147542P.
PR 05-AUG-1999; 99US-0147547P.
PR 05-AUG-1999; 99US-0147549P.
PR 05-AUG-1999; 99US-0147824P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Hodgson DM, Lincoln SE, Russo FD, Spiro PA, Banville SC;
PI Bratcher SR, Dufour GE, Cohen HJ, Rosen BH, Hillman JL;
PI Jones AL, Yu JY, Greenawalt LB, Panzer SR, Roseberry AM, Wright RJ;
PI Daniels SE;
XX
XX WPI; 2001-016511/02.
XX
PT Fifty two human polynucleotides, referred to as DITHP polynucleotides,
PT useful in the diagnosis and treatment of cancer, immune disorders and
PT neurological diseases.
XX
XX Claim 1; Page 261-262; 263pp; English.
XX
XX The present sequence is one of fifty-two human polynucleotides for
XX diagnostics and therapeutics. The polynucleotides are referred to as
XX DITHP polynucleotides. The DITHP polynucleotides and polypeptides are
XX useful for diagnosing and treating diseases such as cancer, particularly
XX breast and ovarian cancer, and other cancers of the adrenal gland, bone,
XX bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
XX immune disorders such as Addison's disease, allergies, autoimmune
XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
XX cardiovascular disorders such as myocardial ischaemias; neurological
XX diseases such as cerebral palsy and epilepsy, infectious diseases such as
XX viral, bacterial, fungal and parasitic infections; and endocrine
XX disorders (e.g. disorders of hypothalamus), disorders associated with
XX hypothyroidism and hyperthyroidism, pancreatic disorders (e.g. diabetes
XX mellitus) and metabolic disorders. The DITHP polynucleotides are useful
XX for screening for molecules that bind to or are bound by the encoded
XX polypeptides. The anti-DITHP antibodies are useful in diagnostic assays
XX

SQ	Sequence	2757 BP; 735 A; 622 C; 682 G; 713 T; 0 U; 5 Other;
	Query Match	16.9%; Score 418.2; DB 4; Length 2757;
	Best Local Similarity	56.9%; Pred. No. 3.2e-66;
	Matches	882; Conservative 0; Mismatches 653; Indels 15; Gaps 6;
QY	459 GAGAGAGCGGTGCGCCCTGGAGCTAGCTACGTCAACTCCAACTTGAGCTGCTGCTGAAGA	518
Db	1 GCGGAGCAAGTGGCGCTGGAGCTTGGCTTCACTCAGACCTCGCATCTCAAGA	60
QY	519 GGAGCTGAGGAGCTCAGCGGTGGGCTGGACCTCGCCGGCATGGAGCGAGCTGTGAC	578
Db	61 AGAGCTGGAGGGCTGGAACATCTCGTGGCGCTCTATCAGAAACAGAGAGGCAATTAC	120
QY	579 TGTCCCATGATCCCCCTGGGCTGAAGGAGACCAAGGAGCTGAGCTGTCTACACCGCT	638
Db	121 GATTCCTCGTATCTCTGGCTGAAGGAACGAAAGACGTGCGACTTTGCGAGTGTCT	180
QY	639 GAAGGAGCTGATCTCAGTGCACCTTTGGAGAGACGGCGCTCCTACGAGCGCAAAATCAG	698
Db	181 CAAGGATTTTATCCTGGGAACATTACAGTCAAGATGGCTATTATATGAAGATGAATAATGC	240
QY	699 GGAGCTGAGGCGCTCGCGCAGGCCATCGGACCCCGCCAGCGGAATGAGTGGGCTCGA	758
Db	241 AGATCTTATGATCTGAGACAGCTTGTGAGCGCTAGCCGGGATGAGCCGGGTGGA	300
QY	759 GCTGCTCAGCCTATTACAACAGCTGTGCTTCTGATGGCGCTTCTCCTACCCCTGC	818
Db	201 ACTGCTGATGACATCTTCACTCAGCTGGGCTTGTGCGAGTGTGATTTCTTCCCGCCAC	360
QY	819 CAGGAGCTCGGGCTTCTTCCATGCTAGCTACGCTTACTGGGGTCCGGGCCAGGA	878
Db	361 ACGGCAGATGGAGCTCTGTTCACTCGTATGACTCTCTCACCGGGGTTCCGGTCAGCA	420
QY	879 GCGTCCCTGGCTTCGAGAGGGCAGGCTTCTCTTCAACATCGTGTGCTCCACACGA	938
Db	421 GCAGNACTGTGCTGGAGAGGCCAGTGTCTGTTCAACACTGGGGCCCTTACACCA	480
QY	939 GATTGGGGCGGCCAGGACCGCTCTCTGCAACGAGGGTGCCTCGCGCTATGGAGCGCTT	998
Db	481 GATTGGGACCCGGTGTGATGCGCAGACGAGCTGGGCTGGAGAGTGCATAGATGCTT	540
QY	999 CCAGGGCGCTGGGGCTTTCAGCTCTGAGGAGAACTTCTCCATGCGCCGAGGCC	1058
Db	541 TCAGAGAGCCGAGGGGTTTAAATTTACCTGAAAGACACATTTACCCATCTCTCAAGTTA	600
QY	1059 AGACATGAGCTGCGTCCCTCGGCACTGAGCAGCTCATGATGGCCGAGGCCAGGA	1118
Db	601 CGACATGAGCCCTGCCATGCTCAGGTGTCTGTAATGATGCTTGACAAAGCCCAAGA	660
QY	1119 ATGTGTGTTGAGGCTCTTCAACCTGCTCCATGGCCCCCAGAGCTGCTGGCCCA	1178
Db	661 AAGCGTGTGAGAAATCAGCTTCTCTG-----GGATCCGGAATGAATTTCTCATGCT	714
QY	1179 GCTGGCCCTGGCGAGGAGGCGCCAGTGGCAGCAGTACAGGCTAGTGACCGGAC	1238
Db	715 GGTGAAGGTGGCTCAGGAGGCTGTAAAGTGGAGAGGTCTACCAACAGCTACACGAGC	774
QY	1239 CATGCCCCAGCACCCCTCCAGCATGCTGCTCTCTGAGCTGCGCTGGTGCATGT	1298
Db	775 CATGAGCCAGGCGCGGTGAAGATTAACATCCCTACTCTCTGGGCGAGTTAGCTGGT	834
QY	1299 CAAGCCGAGTATTTCGCTCCCTGGCCCACTACCACTAGCCATGGCCCTCTGGCAGG	1358
Db	835 GAAGGCCCACTACGCGGCCCTGGCCCACTTCTCACTGCACTCTCTCATNGACA	894
QY	1359 CTCCCCAGGACCGAGGAGAGCTCCCAACGACAGCAGGCTTCTCTGAGCCGCC	1414
Db	895 CCAGGTGAAGCCAGGCA CGGATNTGGACCAACGAGGAGAGTGCCTGTCCCACTACGA	954
QY	1415 CCACCTCTCTAAGCCCGAGGCGCTGTGCTGCC--GCAGGAGCTGGAGGCGCAGGA	1472
Db	955 CCACATGCGAGGGGCTGACACCTTGGCCCACTGAAGATGATCAGCAGCGCCGACA	1014
QY	1473 GCTTGGCAAGGCACACCTGAAGCGTGCATCTCTGGGCGAGGAGGCGCTCGCGGTGCA	1532
Db	1015 GCTGGGAAGTCCCACCTTTGCGCAGAGCCATGGCTCATCANGAGGAGTGGTTCGGGAGGC	1074
QY	1533 CGCCCTGTGCGCGCTCTGCGGAGGTGGA CTTGCTTGGGGCTGTGATCTCCACACGCT	1592
Db	1075 GAGCCTCTGCAAGAGAGCTCCG--GANATTGAGGTGCTACAGAGGTGTGNGTCCG--CACA	1132
QY	1593 GCAGCGCTCACTGGCCAAAGTATGGAGCTGCACCGTGGAGGATGACTTCTGTGAGGCTGC	1652
Db	1133 GGAACGCTCCCGCTCAGCTACGTCAGCCAGCA CAGAGGAGGATGACCTGCTGAACTGAT	1192
QY	1653 CGAGCCCCGAGCATCCAGCCTTAAGACCCAC CAGAAAGCCAGAGGCCAGGATGCCACGCT	1712
Db	1193 CGAGCCCCCAGTGTGTTGCTTAAACTGAGCA GAGGTTGACATTATATTGCC--CCATT	1251
QY	1713 GTCCAGGGGAAGGGCCCTGACATCTTCCATCGGCTGGGGCCCTGTCTGTGTTCTCAGC	1772
Db	1252 CTCGAAGCTGACAGTCAACGGA CTTCTTCCAGAAGCTGGGGCCCCCTTATCTGTGTTTCGGC	1311
QY	1773 CAAGAACCGGTGGCGCTGGTGGGGCCGTC CACCTGACCCGAGGAGGGCGGCTTTGG	1832
Db	1312 TAAAGAGCGGTGGACGCTCTCTCGAAGCAT CGCTTCACTGCAAGAGAGGGAATTGG	1371
QY	1833 CTTACGCTTCGGGAGACTCGCTGTCTCTCAT CGCTGCGCTCATTTCCAGGAGCCAGGC	1892
Db	1372 GTTACCTTGAGAGGGAACGCCCGTT CAGGTTCACTTCTGGATCTTACTGCTCTGC	1431
QY	1893 CGCGCGCTGGCTGGAAGGAGGGGCACTA CATTGTGTGATGGGAGCCATGCG	1952
Db	1432 CTCGGTGGCAGGAGCCCGGAAGGAGATTAT ATTTGCTCTCCATTCCAGCTTGTGGATTGTA	1491
QY	1953 GTGGTGGAGACACGCGGAGGTGGTGA CGGAGCTGAAGGCTCGCGAGAGG	2002
Db	1492 GTGGCTGAGCTGAGTGAGGTTATGAAGCT GCTGAGAGCTTTGGCGAGG	1541
RESULT	15	
ABK92255		
ID	ABK92255	standard; DNA; 1671 BP.
XX	AC	ABK92255;
XX	DT	15-AUG-2002 (first entry)
XX	DE	Prostate cancer-associated DNA sequence #141.
XX	KW	Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;
XX	KW	gene therapy; gene; ds.
XX	OS	Mammalia.
XX	FN	WO200230268-B2.
XX	PD	18-APR-2002.
XX	PF	12-OCT-2001; 2001WO-US032045.
XX	PR	13-OCT-2000; 2000US-00687576.
XX	PR	08-DEC-2000; 2000US-00733288.
XX	PR	24-JAN-2001; 2001US-0263957P.
XX	PR	16-MAR-2001; 2001US-0276791P.
XX	PR	06-MAR-2001; 2001US-0276888P.
XX	PR	06-APR-2001; 2001US-0281922P.
XX	PR	24-APR-2001; 2001US-0286214P.
XX	PR	30-APR-2001; 2001US-00847046.
XX	PR	04-MAY-2001; 2001US-0288589P.
XX	PA	(EOSB-) EOS BIOTECHNOLOGY INC.
XX	PI	Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;

XX WPI: 2002-471335/50.
DR P-PSDB; ABG61936.
XX
PT Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.
XX
XX Claim 22; Page 417-418; 436pp; English.
XX
CC The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with prostate cancer-
CC associated polynucleotides (designated PC genes) that selectively
CC hybridize to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
CC sequences
XX
SQ Sequence 1671 BP; 402 A; 459 C; 457 G; 353 T; 0 U; 0 Other;
Query Match 15.4%; Score 379.2; DB 6; Length 1671;
Best Local Similarity 56.0%; Pred. No. 3.3e-59;
Matches 783; Conservative 0; Mismatches 603; Indels 12; Gaps 3;
643 GAGCTGATCTCAGTGCACTTTGAGAGGACGGCGCTCTTACGAGGAGAAATGAGGAG 702
Db 1 GATTTTATCTCGAAACATTACAGTGAAGATGGCTATTTATATGAAGATGAAATGCGAGAT 60
703 CTGAGGCGCTGGCGAGGCGCATCGGACCCCGAGCGGAATGATCGGCGCTGGAGCTG 762
Db 61 CTTATGATCTGAGACAAGCTTGTGCGACGCCCTAGCGCGGATGAGGCGGGTGGAACTG 120
763 CTCACAGCTTATTAACAACAGCTGTGCTTCTCGATGCGCGCTTCTTCAACCCCTGCCAGG 822
Db 121 CTGATGACATACATTCATCAGCTGGGCTTTGTCGAGAGTGCATTCTTCCGCGCCACACGG 180
823 AGCCTGGGCTCTTCTTCACTGGTAGAGTCTGCTTCTGAGTCTGGGGTCCGCGCCAGCAGCT 882
Db 181 CAGATGGGACTCTCTGTTCACTGGTATGACTCTCTCACCGGGGTTCGGGTTCAGCCAGCAG 240
883 GCCTTGGCTTCGAGAAAGGCGAGCTTCTTCTCAACATCGGTGCGCTTCCACACGCGAGATT 942
Db 241 AACCTGTGCTGGAGAAAGCCAGTGTCTCTTTCAACACTGGGCGCTCTTACACCCAGATT 300
943 GGGCGCGCGCCAGGACCGCTCTCTGCAACGAGGGTGC CGCGCGCTATGAGGGCTTCCAG 1002
Db 301 GGGACCCGGTGTGATCGGCAGACGCGAGCTGGGCTGGAGAGTGCCATAGATGCTTTCAG 360
1003 AGGGCGCTGGGCGCTTTCAGCTTCTGAGGAGAACTTCTTCCATGCGCGCGAGCCAGCAGAC 1062
Db 361 AGAGCGCGCAGGGGTTTTAAATTAACCTGAAAGACACATTTTACCCATATCTCCAAAGTTAGCAG 420
1063 ATGAGCGCTGCTGCTCTGCGCACTGGAGCAGCTCATGATGCGCGAGGCGCCAGGAAATGT 1122
Db 421 ATGAGCGCTGCCATGCTCAGCGTGTCTGTCAAATATGATGCTTGCACAGCCCGAGAAAGC 480
1123 GTGTTTGAAGGCTCTCAACACTGCTCTCATGCGCCGCCCAAGACTGCTTGGCCCCAGCTG 1182
Db 481 GTGTTTGAAGAAATCAGCCTTCTCTG-----GGATCGGGAATGAATTTCTTCATGCTGGTG 534
1183 CGCTTGGCGAGGCGCGCCAGCTGGAGCGGAGTACAGGCTAGTGTGACCGGACCCATG 1242
Db 535 AAGGTGGCTCAGGAGGCTGCTAAGGTGGGAGAGGTCTTACCAACAGCAGCTACACGCGCCATG 594

QY 1243 GCCAGGCCACCCCTCCACGACTAGCTGCTGCTCTCTGGAAGTCCCTGGTGCATGTCTCAAG 1302
Db 595 AGCCAGCGCGCGGTGAAGAGAAACATCCCTACTCTCTGGGCCAGCTTAGCCTTGGTGAAG 654
QY 1303 GCCAGTACTTTCGCTCCCTGGGCCACTACAGTAGCAGCATGCGCCCTCTCTCGACGGCTCC 1362
Db 655 GCCACCACTACGCGGCGCTGGCCCACTATTCTCACTGCCATCTCTCTCATCGACCCACG 714
QY 1363 CACGACCGAGGAGAGCTCCACGACGAGAGGTCTTCTCTGAGC-----CCCCCAG 1418
Db 715 GTCAAGCCAGGCAAGGATCTGGACCAACAGGAGAGTGCCTGTCTCCAGCTCTTAGGACCA 774
QY 1419 CTCTCTTAAGCCCGGAGGCGCTCTGCT--GCCGACGAGAGCTGAGGAGCCGAGCAGCTT 1476
Db 775 ATCCAGAGGGGTGACACCCCTTGGCCACTGAAGAAATGATCAGACGCCGACAGCTG 834
QY 1477 GGCAAGGCACACTGAAGCGTGCATCTCTGGGSCAGGAGGCGCTGCGGCTGCAGCC 1536
Db 835 GGGAAAGTCCACATTGCGCAGAGCCATGGCTCATCACGAGGAGTGGTGGGAGGCCAGC 894
QY 1537 CTGTGCGGCTCTCTGCGAGGTTGACCTGCTTGGGCTGTGATCTCCAGAGCTGTCAG 1596
Db 895 CTCTGCAAGAAGCTGCGGAGCAATTGAGGTGCTACAGAAAGTGTGTGTGCCGACAGGAA 954
QY 1597 CGCTCACTGSCCAAGTATGCGGAGCTCGACCGTGAGATGACATTCTGTGAGGCTGCCGAG 1656
Db 955 CGCTCCGCGCTACGTACGCCCCAGCACGAGGAGAGATGACCTGTGAACCTGATGAC 1014
QY 1657 GCCCGGACATCCAGCTTAAGACCCACAGAACCCAGAGCCAGGATGCCAGCCTGTCC 1716
Db 1015 GCCCCAGTGTGTTGCTTAAACTGAGCAAGAGGTTGACATTATATTGCCCCAGTTCTCC 1074
QY 1717 CAGGGAAGGGGCTGACATCTTCATCGGCTGGGCGCTCTGTGTGTCTCAGCCAG 1776
Db 1075 AAGCTGACAGTCAAGGACTTCTTCCAGAAAGCTGGGCGCTTATCTGTGTCTTCCGGCTAAC 1134
QY 1777 AACCGTGGCGGCTGTGTGGGCGCGCTCCACTGACCCGAGAGAGGCGGCTTTGGGCTC 1836
Db 1135 AAGCGTGGAGCGCTCTCTGAGCATCCGTTTCACTGACAGAAAGGGGACTTTGGGGTTC 1194
QY 1837 AGCTTTCGGGAGACTGCGCTGTCTCATGCTGCGCTGCATTCAGGGAGCCAGGCGCG 1896
Db 1195 ACCTTGAGAGGGAACGCGCCCGTTCAAGTTCACTTCTGGAATCCTTACTGCTCTGCTCG 1254
QY 1897 GCGGCTGGCTGAAGAGGCGGACTACATGTGTGCTGAGTGAATGGGAGCCATGCAAGTGG 1956
Db 1255 GTGGCAGGAGCCCGGGAAGGAGATTATTTGTCTCCATTGCTGAGTTGTGAGTTGTAAGTGG 1314
QY 1957 TGGAGACACGCGGAGGTGTGACGAGCTGAAGGCTGCGGAGAGGCGGCGGCGCAGCCTG 2016
Db 1315 CTGACGCTGAGTGAGGTATGAGAGCTGTGAAGAGCTTTGGCGGAGGACGAGATGAGATG 1374
QY 2017 CAGGTGGTGTGCTGTG 2034
Db 1375 AAAGTGTGAGCCTCTG 1392

Search completed: July 13, 2004, 02:51:03
Job time : 645.213 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2004, 20:08:35 ; Search time 6334.33 Seconds
(without alignments)
16894.276 Million cell updates/sec

Title: US-10-697-266-1
Perfect score: 2469
Sequence: 1 tcggcgcccggtggtggg.....aaaaaaaaaaaaaaaaaaaaa 2469

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues 6940544
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.ov.*

22: em.pat.*

23: em.ph.*

24: em.pl.*

25: em.ro.*

26: em.sts.*

27: em.un.*

28: em.vi.*

29: em.htg_hum.*

30: em.htg_inv.*

31: em.htg_other.*

32: em.htg_mus.*

33: em.htg_pln.*

34: em.htg_rdt.*

35: em.htg_rdt.*

36: em.htg_rdt.*

37: em.htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2469	100.0	2469	6	AR269115	Sequence
2	2469	100.0	2469	6	AX642309	Sequence
3	2105	85.3	2837	6	AX675063	Sequence
4	2105	85.3	3763	9	BC025767	Homo sapi
5	2019	81.8	3647	9	AB067516	Homo sapi
6	1869	75.7	2088	9	AY082588	Homo sapi
7	1102.2	44.6	1932	10	MMU43194	Mus musculus
8	1043.8	42.3	2301	10	BC052010	Mus musculus
9	952.8	38.6	4729	6	BD183400	Novel gen
10	558.8	22.6	3230	4	CFA347749	Canis fam
11	524.6	21.2	764	6	BD214873	Novel hum
12	496.4	20.1	3516	9	BC036447	Homo sapi
13	494.8	20.0	2061	9	HSA347750	Homo sapi
14	494.8	20.0	3500	6	AX835273	Sequence
15	494.8	20.0	3500	9	AK098246	Homo sapi
16	487.8	19.8	3019	6	AX077672	Sequence
17	485.6	19.7	3397	9	HSB803254	Homo sapi
18	477.4	19.3	2233	5	BC044556	Danio rer
19	475.8	19.3	2224	5	AY391456	Danio rer
20	443.4	18.0	2262	9	AF268032	Homo sapi
21	418.2	16.9	2757	6	AX591072	Sequence
22	368	14.9	19025	6	AR269116	Sequence
23	368	14.9	19025	6	AX642311	Sequence
24	368	14.9	179057	2	AF276758	Homo sapi
25	368	14.9	179941	9	AC105118	Homo sapi
26	351.2	14.2	178343	2	AC024049	Homo sapi
27	311	12.6	3798	9	AK126506	Homo sapi
28	272.8	11.0	2403	3	AF132025	Drosophil
29	168.2	6.8	61188	2	AC135076	Homo sapi
30	165.4	6.7	123016	2	AF189001	Homo sapi
31	165.4	6.7	124060	9	AC087793	Homo sapi
32	165	6.7	165	6	BD213706	Novel hum
33	159.8	6.5	179057	2	AF276758	Homo sapi
34	149.6	6.1	2300	6	AX833345	Sequence
35	149.6	6.1	2300	9	AK095001	Homo sapi
36	146.6	5.9	232687	2	AC116393	Mus muscu
37	143.4	5.8	261608	2	AC097158	Rattus no
38	143.4	5.8	272301	2	AC133673	Rattus no
39	127	5.1	3174	2	AC014171	Drosophil
40	127	5.1	185200	3	AC092401	Drosophil
41	127	5.1	191558	3	AC013431	Drosophil
42	127	5.1	303345	3	AE003500	Drosophil
43	120.2	4.9	511	9	HSA323305	Homo sapi
44	110	4.5	5234	9	AB025194	Homo sapi
45	108.4	4.4	2925	6	AX713572	Sequence

ALIGNMENTS

RESULT 1	AR269115	AR269115	Sequence 1	2469 bp	DNA	linear	PAT 10-APR-2003
LOCUS	AR269115	Sequence 1	from patent	US 6500655.			
DEFINITION	AR269115	Sequence 1	from patent	US 6500655.			
ACCESSION	AR269115	Sequence 1	from patent	US 6500655.			
VERSION	AR269115.1	GI:29699963					
KEYWORDS	Unknown.						
SOURCE	Unknown.						
ORGANISM	Unknown.						
REFERENCE	1 (bases 1 to 2469)						
AUTHORS	Rusch, D., Ketchum, K.A., Di Francesco, V. and Beasley, E.M.						
TITLE	Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof						
JOURNAL	Patent: US 6500655-A 1 31-DEC-2002;						

FEATURES		Location/Qualifiers	
source		1..2469	
		/organism="unknown"	
		/mol_type="genomic DNA"	
ORIGIN			
Query Match		100.0%;	Score 2469; DB 6; Length 2469;
Best Local Similarity		100.0%;	Pred. No. 0;
Matches 2469;		Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	TCGCGGCCACAGGTGCTGCGGCGGCCCTTAGCCCGGCTGCGAGCGCTGCGCGAGCGCGG	60
Db	1	TCGCGGCCACAGGTGCTGCGGCGGCCCTTAGCCCGGCTGCGAGCGCTGCGCGAGCGCGG	60
Qy	61	GCTGGCTGACCCCGAGGAGACCCCGACGCGAGCGGGTGCAGCATGATCTCGAGGAGAGG	120
Db	61	GCTGGCTGACCCCGAGGAGACCCCGACGCGAGCGGGTGCAGCATGATCTCGAGGAGAGG	120
Qy	121	CGGAGCGGCGGCGCGCGGAGGAGAGCCCGCGCTGCAGATATCTAGGAGAAACCC	180
Db	121	CGGAGCGGCGGCGCGGAGGAGAGCCCGCGCTGCAGATATCTAGGAGAAACCC	180
Qy	181	AGGAAACACGTGTGAGCTCTTTACGGGGAGAGACGGGAAGCCCTGAGAGACGTGTGTGG	240
Db	181	AGGAAACACGTGTGAGCTCTTTACGGGGAGAGACGGGAAGCCCTGAGAGACGTGTGTGG	240
Qy	241	TGAGAGGGTGTGCGGTTCACAGAGGGGAAGACCGAGTGTGTGACGTTGCGCCCATG	300
Db	241	TGAGAGGGTGTGCGGTTCACAGAGGGGAAGACCGAGTGTGTGACGTTGCGCCCATG	300
Qy	301	AATCCGCGAGCTTCATGACGTGGCTGTGACTCCCTGACGAGATCCAGTGCAGCTG	360
Db	301	AATCCGCGAGCTTCATGACGTGGCTGTGACTCCCTGACGAGATCCAGTGCAGCTG	360
Qy	361	CAGAGCGGAGGCCCCAGATTCCACAGAGATTGACAAAGAGCTGACAGATGCGGACGGC	420
Db	361	CAGAGCGGAGGCCCCAGATTCCACAGAGATTGACAAAGAGCTGACAGATGCGGACGGC	420
Qy	421	GCTGAGAACCTTACAGAGCCACAGCAACACCGGGTGAAGAGAGCGTTCGCCCTGGAG	480
Db	421	GCTGAGAACCTTACAGAGCCACAGCAACACCGGGTGAAGAGAGCGTTCGCCCTGGAG	480
Qy	481	CTGAGCTACGCTCAACTCCAACTGACGCTGTGAAGGAGAGCTGAGGAGAGCTCAGCGGT	540
Db	481	CTGAGCTACGCTCAACTCCAACTGACGCTGTGAAGGAGAGCTGAGGAGAGCTCAGCGGT	540
Qy	541	GGCTGTGACCTGCGCGGCATGGAGCGAAGCTGTCACTGTCCCATATGATCCCGCTGGC	600
Db	541	GGCTGTGACCTGCGCGGCATGGAGCGAAGCTGTCACTGTCCCATATGATCCCGCTGGC	600
Qy	601	CTGAAGGAGACCAAGGAGCTGGACTGGTCTACACCGCTGAAGGAGCTGATCTCAGTGCAC	660
Db	601	CTGAAGGAGACCAAGGAGCTGGACTGGTCTACACCGCTGAAGGAGCTGATCTCAGTGCAC	660
Qy	661	TTTGGAGAGGACCGCGCTCTACAGGCGAAGATCAGGGAGCTGAGGAGCCCTGCGGCAG	720
Db	661	TTTGGAGAGGACCGCGCTCTACAGGCGAAGATCAGGGAGCTGAGGAGCCCTGCGGCAG	720
Qy	721	GCCATGCGGACCCCGACCGGATGAGTGGGCGCTGAGAGCTGTCTACAGCCATTATCAAC	780
Db	721	GCCATGCGGACCCCGACCGGATGAGTGGGCGCTGAGAGCTGTCTACAGCCATTATCAAC	780
Qy	781	CAGCTGTGCTTCCTGAGTGGCGCTTCTCACCCCTGCCAGGAGCTCTGCTCTTCTTC	840
Db	781	CAGCTGTGCTTCCTGAGTGGCGCTTCTCACCCCTGCCAGGAGCTCTGCTCTTCTTC	840
Qy	841	CACGTGTACGACTCGCTTACTTGGGGTCCCGGCCAGCAGCTGCTTGGCTTCGAGAAG	900
Db	841	CACGTGTACGACTCGCTTACTTGGGGTCCCGGCCAGCAGCTGCTTGGCTTCGAGAAG	900
Qy	901	GGCAGCGCTTCTTCAACATCGGTGCGCTTCACACGAGATTGGGGCGGCCAGGACCGC	960
Db	901	GGCAGCGCTTCTTCAACATCGGTGCGCTTCACACGAGATTGGGGCGGCCAGGACCGC	960
Qy	961	TCCTGCAACGAGGGTGCCTGCGCGCGCTATGAGAGCGCTTCCAGAGGCGCGCTGGGCGCTTC	1020
Db	961	TCCTGCAACGAGGGTGCCTGCGCGCGCTATGAGAGCGCTTCCAGAGGCGCGCTGGGCGCTTC	1020
Qy	1021	AGCCTCTGAGGAGAACTTCTCCCATGCGCCGAGCCACAGACATGAGCGCTGGCTCCCTC	1080
Db	1021	AGCCTCTGAGGAGAACTTCTCCCATGCGCCGAGCCACAGACATGAGCGCTGGCTCCCTC	1080
Qy	1081	TGGGCACTGAGGAGAGCTCATGATGGCCCAAGGATGTGTGTTGAGGGCTCTCA	1140
Db	1081	TGGGCACTGAGGAGAGCTCATGATGGCCCAAGGATGTGTGTTGAGGGCTCTCTCA	1140
Qy	1141	CCACCTGCTTCCATGCGCCCCCAAGACTGCTGCGCCAGTGCCTGGCGCAGGAGGCC	1200
Db	1141	CCACCTGCTTCCATGCGCCCCCAAGACTGCTGCGCCAGTGCCTGGCGCAGGAGGCC	1200
Qy	1201	GCCAGGTGCGACGCGAGTACAGGCTAGTGACACCGAACCATGGCCACGACCCGCTCCAC	1260
Db	1201	GCCAGGTGCGACGCGAGTACAGGCTAGTGACACCGAACCATGGCCACGACCCGCTCCAC	1260
Qy	1261	GACTAGTGGCTGTCTCTGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1320
Db	1261	GACTAGTGGCTGTCTCTGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1320
Qy	1321	CTGGCCCACTTACCACTGAGCCATGCGCGCTCTGCGAGCGCTCCCGACGACCGAGGAGAG	1380
Db	1321	CTGGCCCACTTACCACTGAGCCATGCGCGCTCTGCGAGCGCTCCCGACGACCGAGGAGAG	1380
Qy	1381	CTCCCCACGCAAGAGAGGTCTTCTGCGACCCCGCCCACTCTTAAAGCCCGAGGCGCT	1440
Db	1381	CTCCCCACGCAAGAGAGGTCTTCTGCGACCCCGCCCACTCTTAAAGCCCGAGGCGCT	1440
Qy	1441	GTGCTGCGCAGAGCTTGGAGGAGCGCAGGAGCTTGGAGGAGCAGACCTGAAGCGGTC	1500
Db	1441	GTGCTGCGCAGAGCTTGGAGGAGCGCAGGAGCTTGGAGGAGCAGACCTGAAGCGGTC	1500
Qy	1501	ATCTGCGGCGAGGAGGCGCTGCGGCTGCGAGCGCTGTCGCGCGCTCTGCGCGAGGTG	1560
Db	1501	ATCTGCGGCGAGGAGGAGGCGCTGCGGCTGCGAGCGCTGTCGCGCGCTCTGCGCGAGGTG	1560
Qy	1561	GACTGCTTCCGGCTGTGATCTCCAGACGCTGCGAGCGCTCACTGGGCAAGTATGCGGAG	1620
Db	1561	GACTGCTTCCGGCTGTGATCTCCAGACGCTGCGAGCGCTCACTGGGCAAGTATGCGGAG	1620
Qy	1621	CTGACCGCTGAGATGATCTGCTGAGGCTGCGAGGCGCCCGGACATCCAGGCTTAAGAC	1680
Db	1621	CTGACCGCTGAGATGATCTGCTGAGGCTGCGAGGCGCCCGGACATCCAGGCTTAAGAC	1680
Qy	1681	CACCAAGGCGAGGCGCAGGATGCGACGCTGTCTCCAGGGGAGGGGCGCTGACATCTTC	1740
Db	1681	CACCAAGGCGAGGCGCAGGATGCGACGCTGTCTCCAGGGGAGGGGCGCTGACATCTTC	1740
Qy	1741	CATCGGCTGGGCGCTGTCTGTGTTCTAGCCAAAGAACCGGTGCGCGCTGCTGGGCGCC	1800
Db	1741	CATCGGCTGGGCGCTGTCTGTGTTCTAGCCAAAGAACCGGTGCGCGCTGCTGGGCGCC	1800
Qy	1801	GTCCACCTGACCGAGGAGGCGGCTTTGGCTTCTACGCTTCCGGGAGAGCTCGCTGTGTC	1860
Db	1801	GTCCACCTGACCGAGGAGGCGGCTTTGGCTTCTACGCTTCCGGGAGAGCTCGCTGTGTC	1860
Qy	1861	CTCATGCTGCTGCTTCTTCCAGGGAGCCAGGCGCGCGCTGCTGCTGAAGGAGGCGAC	1920
Db	1861	CTCATGCTGCTGCTTCTTCCAGGGAGCCAGGCGCGCGCTGCTGCTGAAGGAGGCGAC	1920
Qy	1921	TACATGTGTGAGTGAATGGGAGCCATGAGTGTGGAGACACCGGAGGAGTGTGACG	1980
Db	1921	TACATGTGTGAGTGAATGGGAGCCATGAGTGTGGAGACACCGGAGGAGTGTGACG	1980
Qy	1981	GAGCTGAAGGCTGCGGAGAGGCGGCGCCAGCTGCGAGTGTGCTGCTGCTGCCAGC	2040
Db	1981	GAGCTGAAGGCTGCGGAGAGGCGGCGCCAGCTGCGAGTGTGCTGCTGCTGCCAGC	2040

QY	2041	TCTAGACTGCCAGCTTGGGGGACCGCGGCCGCTCTGCTGGGCCCAAGGGGGCTTCTA	2100	
Db	2041	TCTAGACTGCCAGCTTGGGGGACCGCGGCCGCTCTGCTGGGCCCAAGGGGGCTTCTA	2100	
QY	2101	AGGAGCCAGAGGAGCATGTTTGCAGAGACCCCGGCATCCACGTGGGCCAGTCCCGGGCC	2160	
Db	2101	AGGAGCCAGAGGAGCATGTTTGCAGAGACCCCGGCATCCACGTGGGCCAGTCCCGGGCC	2160	
QY	2161	CTCCTCAACTGGAGCCGAAAGGCCCAAGGAGGCAAGACTGGAGGCTGCCCGCAGCCCTGT	2220	
Db	2161	CTCCTCAACTGGAGCCGAAAGGCCCAAGGAGGCAAGACTGGAGGCTGCCCGCAGCCCTGT	2220	
QY	2221	GCCCGAGTGAAGCAGCTCCGGCCCTCATCTTGAAGCAACCGGTGGCCGTGAGGGCA	2280	
Db	2221	GCCCGAGTGAAGCAGCTCCGGCCCTCATCTTGAAGCAACCGGTGGCCGTGAGGGCA	2280	
QY	2281	GGATCCCTGCAGCCCTCAGCCCTGGCTCCAGCTGGCAGCAAGCAGCATGCCCTCC	2340	
Db	2281	GGATCCCTGCAGCCCTCAGCCCTGGCTCCAGCTGGCAGCAAGCAGCATGCCCTCC	2340	
QY	2341	CCACCAGAGGACCTCCGGGCAATGCTGTCCGGCTCATGTGGAGGCTGCTCGGGCA	2400	
Db	2341	CCACCAGAGGACCTCCGGGCAATGCTGTCCGGCTCATGTGGAGGCTGCTCGGGCA	2400	
QY	2401	CCTGCTGCCCTTAAAGACTGGTCAGACTGTCTGAAGAAAAAAGAAAAA	2460	
Db	2401	CCTGCTGCCCTTAAAGACTGGTCAGACTGTCTGAAGAAAAAAGAAAAA	2460	
QY	2461	AAAAAAAAA 2469		
Db	2461	AAAAAAAAA 2469		
RESULT 2				
LOCUS	AX642309	2469 bp	DNA	linear PAT 21-FEB-2003
DEFINITION	Sequence 1 from Patent WO02061062.			
ACCESSION	AX642309			
VERSION	AX642309.1 GI:28474733			
KEYWORDS	.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1 Rusch, D., Ketchum, K., di Francesco, V. and Beasley, E.M.			
TITLE	Human protein kinase n-like polypeptide and uses thereof			
JOURNAL	Patent: WO 02061062-A 1 08-AUG-2002;			
PE CORPORATION (NY) (US)				
FEATURES	Location/Qualifiers			
source	1..2469			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			
ORIGIN				
Query Match	100.0%; Score 2469; DB 6; Length 2469;			
Best Local Similarity	100.0%; Pred. No. 0;			
Matches 2469;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	TGCGGGCCAGGTGGTGGGGGGCCCTAGCCCGGCTGGGAGCGCTGGCGAGCGGCGG	60	
Db	1	TGCGGGCCAGGTGGTGGGGGGCCCTAGCCCGGCTGGGAGCGCTGGCGAGCGGCGG	60	
QY	61	GCTGGCTGACCCCGAGGGACCCCGAGCGAGCGGGTGGCGGATGATCTGTGAGGAGAGG	120	
Db	61	GCTGGCTGACCCCGAGGGACCCCGAGCGAGCGGGTGGCGGATGATCTGTGAGGAGAGG	120	
QY	121	CGGAGCGCGGGCGCGCGGAGGAGCGCCGGCTGCAGATATCTAGGAGAAAAACC	180	
Db	121	CGGAGCGCGGGCGCGCGGAGGAGCGCCGGCTGCAGATATCTAGGAGAAAAACC	180	
QY	181	AGGAAACACGTGTGAGCTTTTACGGGGAACAGCGGAAGGCGCTGAGAGACGTGTGGCG	240	

Db	181	AGGAAACACGTGTGAGCTTTTACGGGGAACAGCGGGAAGGCGCTGAGAGACGTGTGGG	240	
QY	241	TGGAGAGGGTGTGGGTCCACAGAGGGGAAGAACCAAGTGCCTGTGACGCTTGGCCCCCATG	300	
Db	241	TGGAGAGGGTGTGGGTCCACAGAGGGGAAGAACCAAGTGCCTGTGACGCTTGGCCCCCATG	300	
QY	301	AATCCGCAAGCTTCAATGCAAGTGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGCCAGCTG	360	
Db	301	AATCCGCAAGCTTCAATGCAAGTGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGCCAGCTG	360	
QY	361	CAGAGCCGAGGGCCAGATTCACAGCAGATTTGACAGAGAGCTGCAGATGCGGAGCGGC	420	
Db	361	CAGAGCCGAGGGCCAGATTCACAGCAGATTTGACAGAGAGCTGCAGATGCGGAGCGGC	420	
QY	421	GCTGAGAACTCTACAGAGCCACAGCAACCAACCGGCTGAGAGAGCGGTGGCCCTGGAG	480	
Db	421	GCTGAGAACTCTACAGAGCCACAGCAACCAACCGGCTGAGAGAGCGGTGGCCCTGGAG	480	
QY	481	CTGAGCTACGTCAACTCCAACTGCAAGTGTCTGAAGGAGAGCTGGAGAGCTCAGCGGT	540	
Db	481	CTGAGCTACGTCAACTCCAACTGCAAGTGTCTGAAGGAGAGCTGGAGAGCTCAGCGGT	540	
QY	541	GGCGTGGAGCCCTGGCGCGCATGGGAGCGAGCTGCTACCTCCCATGATCCCTGGGC	600	
Db	541	GGCGTGGAGCCCTGGCGCGCATGGGAGCGAGCTGCTACCTCCCATGATCCCTGGGC	600	
QY	601	CTGAAGAGAGCAAGAGAGCTGGAGCTTCAACCGCTGAAGAGAGCTGATCTCAGTGAC	660	
Db	601	CTGAAGAGAGCAAGAGAGCTGGAGCTTCAACCGCTGAAGAGAGCTGATCTCAGTGAC	660	
QY	661	TTTGGAGAGAGCGGCGCTCTCAAGAGCAGAAATCAGGAGAGCTGGAGAGCTCGGGCAG	720	
Db	661	TTTGGAGAGAGCGGCGCTCTCAAGAGCAGAAATCAGGAGAGCTGGAGAGCTCGGGCAG	720	
QY	721	GCCATGCGGAGCCCGAGCGGAGTGGAGTGGGGCTGGAGAGCTGCTCAGAGCTATTACAC	780	
Db	721	GCCATGCGGAGCCCGAGCGGAGTGGAGTGGGGCTGGAGAGCTGCTCAGAGCTATTACAC	780	
QY	781	CAGCTGTGCTTCTCGATGCGCGCTTCTCAACCGCTGCAGAGAGCTCGGGCTCTTCTTC	840	
Db	781	CAGCTGTGCTTCTCGATGCGCGCTTCTCAACCGCTGCAGAGAGCTCGGGCTCTTCTTC	840	
QY	841	CACGTGTAAGTCTGCTTACTGGGGTCCCGGCCAGCAGCGCTGCGCTTCGAGAG	900	
Db	841	CACGTGTAAGTCTGCTTACTGGGGTCCCGGCCAGCAGCGCTGCGCTTCGAGAG	900	
QY	901	GGCAGCGTCTCTTCAACATCGGTGCGCTCCACACGAGAGTTGGGGCGCGGAGACCGC	960	
Db	901	GGCAGCGTCTCTTCAACATCGGTGCGCTCCACACGAGAGTTGGGGCGCGGAGACCGC	960	
QY	961	TCCTGACACGAGGGTGGCGCGCTATGAGGCGCTTCCAGAGGCGCGCTGGGGCTTC	1020	
Db	961	TCCTGACACGAGGGTGGCGCGCTATGAGGCGCTTCCAGAGGCGCGCTGGGGCTTC	1020	
QY	1021	AGCTTCTGAGGAGAACTTCTCCATGCGCGAGCCAGACATGAGCGCTGCGTCCCTC	1080	
Db	1021	AGCTTCTGAGGAGAACTTCTCCATGCGCGAGCCAGACATGAGCGCTGCGTCCCTC	1080	
QY	1081	TGGGCACTGGAGCAGCTCATGATGGCCCGAGCGGAGGAGTGTGTTTGAAGGGCTCTCA	1140	
Db	1081	TGGGCACTGGAGCAGCTCATGATGGCCCGAGCGGAGGAGTGTGTTTGAAGGGCTCTCA	1140	
QY	1141	CCACTGCTCATGGGCGCCCAAGACTGCTGGCGCCAGCTGCGCTGCGCGAGAGGCC	1200	
Db	1141	CCACTGCTCATGGGCGCCCAAGACTGCTGGCGCCAGCTGCGCTGCGCGAGAGGCC	1200	
QY	1201	GCCAGGTGGCAGCGAGTACAGGCTAGTGACCGGACCATGGCCCCAGCCCGCTCCAC	1260	
Db	1201	GCCAGGTGGCAGCGAGTACAGGCTAGTGACCGGACCATGGCCCCAGCCCGCTCCAC	1260	
QY	1261	GACTACGTGCTGTCTCTGAGCTGCTGTCATGTCGAAGGCGGAGTACTTCCGCTCC	1320	

QY 801 GGGCTTCTCAGCCCTGCGAGGCTCGGGCTCTTCTTCCACTGTGACGACTCGCTTAC 860
DB |||||
DB 661 GGGCTTCTCAGCCCTGCGAGGCTCGGGCTCTTCTTCCACTGTGACGACTCGCTTAC 720
QY |||||
DB 861 TGGGGTCCCGGCGCCAGCAGCGTGCCTTTCAGAGAGGCGAGGTTCTCTTCAACAT 920
DB |||||
DB 721 TGGGGTCCCGGCGCCAGCAGCGTGCCTTTCAGAGAGGCGAGGTTCTCTTCAACAT 780
QY |||||
DB 921 CGGTGCTTCCACACGAGATTGGGCGCGCCAGGACCGCTCTCTGCAACGAGGCTGCGCG 980
DB |||||
DB 781 CGGTGCTTCCACACGAGATTGGGCGCGCGCCAGGACCGCTCTCTGCAACGAGGCTGCGCG 840
QY |||||
DB 981 CGCGCTATGAGAGGCTTCCAGAGGCGCGCTGGGGCTTTCAGGCTCTCTGAGGAGAACTT 1040
DB |||||
DB 841 CGCGCTATGAGAGGCTTCCAGAGGCGCGCTGGGGCTTTCAGGCTCTCTGAGGAGAACTT 900
QY |||||
DB 1041 CTCCATGCGCGAGCCACAGACATGAGCGCTGCGTCCCTCTGCGACTGGAGCGACTCAT 1100
DB |||||
DB 901 CTCCATGCGCGAGCCACAGACATGAGCGCTGCGTCCCTCTGCGACTGGAGCGACTCAT 960
QY |||||
DB 1101 GATGCGCCAGGCGCCAGGAATGTGTGTTGAGGGCTCTCACCACTGCTTCCATGGCCCC 1160
DB |||||
DB 961 GATGCGCCAGGCGCCAGGAATGTGTGTTGAGGGCTCTCACCACTGCTTCCATGGCCCC 1020
QY |||||
DB 1161 CCAAGACTGCTGCGCCAGCTGGGCTGGCGCAGAGGCGCGCAGGTGGAGCGAGTA 1220
DB |||||
DB 1021 CCAAGACTGCTGCGCCAGCTGGGCTGGCGCAGAGGCGCGCAGGTGGAGCGAGTA 1080
QY |||||
DB 1221 CAGGCTAGTGACCGGACCATGCGCCAGCCACCGCTCCACGACTACGTGCTGTCTCTG 1280
DB |||||
DB 1081 CAGGCTAGTGACCGGACCATGCGCCAGCCACCGCTCCACGACTACGTGCTGTCTCTG 1140
QY |||||
DB 1281 GACTGCTGTGTGATGCAAGGCGAGTAATTCGCTTCCCTGGGCCACTACAGCTAGC 1340
DB |||||
DB 1141 GACTGCTGTGTGATGCAAGGCGAGTAATTCGCTTCCCTGGGCCACTACAGCTAGC 1200
QY |||||
DB 1341 CATGSCCTCTGCGAGGCTGCCAGCGACCGAGGAGGCTCCCAAGCAGGAGCGAGT 1400
DB |||||
DB 1201 CATGSCCTCTGCGAGGCTGCCAGCGACCGAGGAGGCTCCCAAGCAGGAGCGAGT 1260
QY |||||
DB 1401 CTTCTCTGAGCCCGCCACTCTCTTAAGCCCGAGGCGCTGTGCTGCGCAGGAGCTGGA 1460
DB |||||
DB 1261 CTTCTCTGAGCCCGCCACTCTCTTAAGCCCGAGGCGCTGTGCTGCGCAGGAGCTGGA 1320
QY |||||
DB 1461 GGAGCGAGGAGCTGTGGCAAGGACACCTGAAAGCGTGCATCTCTGGGCGAGGAGGCG 1520
DB |||||
DB 1321 GGAGCGAGGAGCTGTGGCAAGGACACCTGAAAGCGTGCATCTCTGGGCGAGGAGGCG 1380
QY |||||
DB 1521 GCTGGGCTGACGCGCTGTGCGCGTCTGCGCAGGTGGACCTGCTTGGGCTGTGAT 1580
DB |||||
DB 1391 GCTGGGCTGACGCGCTGTGCGCGTCTGCGCAGGTGGACCTGCTTGGGCTGTGAT 1440
QY |||||
DB 1581 CTCCAGAGCGTGTGAGCGCTCACTGGCCAAAGTATGCGGAGCTCGACCGTGGAGTGA 1640
DB |||||
DB 1441 CTCCAGAGCGTGTGAGCGCTCACTGGCCAAAGTATGCGGAGCTCGACCGTGGAGTGA 1500
QY |||||
DB 1641 CTGTGAGGCTGCGAGGCGCGGACATCAGCCTTAAGACCCACAGAGCCAGAGGCGAG 1700
DB |||||
DB 1501 CTGTGAGGCTGCGAGGCGCGGACATCAGCCTTAAGACCCACAGAGCCAGAGGCGAG 1560
QY |||||
DB 1701 GATGCCAGCGCTGTCAGGCGAGGCGCTGACATCTTCCATCGGCTGGGCGCCCTGTC 1760
DB |||||
DB 1561 GATGCCAGCGCTGTCAGGCGAGGCGCTGACATCTTCCATCGGCTGGGCGCCCTGTC 1620
QY |||||
DB 1761 TGTGTTCTCAGCAAGAACCGGTGGCGCTGTGGGCGCGCTGCCACCTGACCCGAGGAGA 1820
DB |||||
DB 1621 TGTGTTCTCAGCAAGAACCGGTGGCGCTGTGGGCGCGCTGCCACCTGACCCGAGGAGA 1680
QY |||||
DB 1821 GGGCGCTTGTGCTTCAAGCTTGGGGAGACTGCGCTCTCTCATCGCTGCGCTCATTTCC 1880
DB |||||
DB 1681 GGGCGCTTGTGCTTCAAGCTTGGGGAGACTGCGCTCTCTCATCGCTGCGCTCATTTCC 1740
QY |||||
DB 1881 AGGAGGCCAGGCGCGGCGCTGGGCTTGAAGAGGCGGAGTACATTTGTGTGAGTGAATGG 1940

DB 1741 AGGAGGCCAGGCGCGGCGCTGGCTGAAGAGGCGGAGTACATTTGTGTGAGTGAATGG 1800
QY |||||
DB 1941 GCAGCCATGCGAGCTGTGAGACACGCGAGAGTGTGACGAGCTGAAGGCTGCGGAGA 2000
DB |||||
DB 1801 GCAGCCATGCGAGCTGTGAGACACGCGAGAGTGTGACGAGCTGAAGGCTGCGGAGA 1860
QY |||||
DB 2001 GCGCGGCGCCAGCTGCGAGGTGTGTGCTGTGCGCCAGCTCTAGACTGCCAGCTTGGG 2060
DB |||||
DB 1861 GCGCGGCGCCAGCTGCGAGGTGTGTGCTGTGCGCCAGCTCTAGACTGCCAGCTTGGG 1920
QY |||||
DB 2061 GSAACGCGCGCCCTGCTGCTGGGCGCCAGGGGCTTTTAAGAGCGAGGAGGAGATGG 2120
DB |||||
DB 1921 GSAACGCGCGCCCTGCTGCTGGGCGCCAGGGGCTTTTAAGAGCGAGGAGGAGATGG 1980
QY |||||
DB 2121 TTGCAAGACCCCGGATCCAGCTGGGCGAGTCCCGGCGCCCTCTCACTGGAGCGGAAA 2180
DB |||||
DB 1981 TTGCAAGACCCCGGATCCAGCTGGGCGAGTCCCGGCGCCCTCTCACTGGAGCGGAAA 2040
QY |||||
DB 2181 GSCCCAGCAGGCGAAGACTGGAGGCTGCCCCAGCGCTGTGCCAGTGAAGCCAGCTCC 2240
DB |||||
DB 2041 GSCCCAGCAGGCGAAGACTGGAGGCTGCCCCAGCGCTGTGCCAGTGAAGCCAGCTCC 2100
QY |||||
DB 2241 GSCCTCATCTTGAAGACCCAGGGTGGCGGTGAGGGCCAGGATCCCTGCAAGCCCTCAG 2300
DB |||||
DB 2101 GSCCTCATCTTGAAGACCCAGGGTGGCGGTGAGGGCCAGGATCCCTGCAAG-CCTCAG 2159
QY |||||
DB 2301 CCTGGCTCCAGCTGGCAGCAGCAGGAGTGGCTGCGGCGCTGCTGCGGCGCTGCTGCCATTAAAGAC 2420
DB |||||
DB 2160 CCTGGCTCCAGCTGGCAGCAGCAGGAGTGGCTGCGGCGCTGCTGCGGCGCTGCTGCCATTAAAGAC 2219
QY |||||
DB 2361 CAATGCTGTCCGCTCATGCTGGAGGTGCTGCTGCGGCGCTGCTGCGGCGCTGCTGCCATTAAAGAC 2420
DB |||||
DB 2220 CAATGCTGTCCGCTCATGCTGGAGGTGCTGCTGCGGCGCTGCTGCGGCGCTGCTGCCATTAAAGAC 2279
QY |||||
DB 2421 TGCTCAGACCTGTCTGA 2437
DB 2280 TGCTCAGACCTGTCTGA 2296

RESULT 4
BC025767 3763 bp mRNA linear PRI 06-OCT-2003
BC025767 Homo sapiens rhophilin, Rho GTPase binding protein 1, mRNA (cdna
clone MGC:34714 IMAGE:5211354), complete cds.
ACCESSION BC025767.1 GI:19343950
VERSION MGC.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3763)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, J., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schaez, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Faney, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinaki, M.J., Skalek, U., Smaluk, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 3763)

Direct Submission

Submitted (06-MAR-2002)

National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapsb@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc_mgc@hri.nih.gov

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,

Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,

Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,

Teurgeson, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAC Plate: 49 Row: j Column: 5
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 19882236.

FEATURES

source

1. .3763

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:34714 IMAGE:5211354"

/tissue_types="Blood, adult leukocytes"

/clone_lib="NIH MGC_118"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. .3763

/gene="RHPN1"

/note="synonyms: KIAA1929, RHPN, ODF5"

/db_xref="LocusID:114822"

102. .2114

/codon_start=1

/product="RHPN1 protein"

/protein_id="AAH25767.1"

/db_xref="GI:19343951"

/db_xref="LocusID:114822"

/translations="MILEERPDPGAGEESPRLOQCDLSLTQICGOLQSRRAQIHOQI
DKLEQFTGAEINLYRATSNRNVRETVALELSYNSNLQLLKELELSGVDPRHGS
EATVPMPLGKETKELDSNTPLEKELISVHFGEDGASYAEARELEALRQAMETPSR
NESLELLTYNQLCFDLPARSLGLFFHWYDSLTFVPAQKALAPFKGSVLIF
NIGALHTQIGARDQSTEGARAMEAFQRAAGASLLRNFENSHAPSDMSAASLCL
EQIMMAQPCVFEGISPPAPQDCIAQLRAQEAQVAAEYRLVHRTPMAQPPVHD
YVPVSWTALVHVHAEYFRSLAHVAMALCGSPATGELPTHEQVFLQPTSSKPRG
PVLPOELEERROIGKALHRAILQGEALRIHALCRVLRVDLILRAVISOTLQSLAK
YALDRDDCEAAEADPIQPKHKEPEABMPILSQKGPDIFHRLGPLSVFSKNPW
RLVGPVHLTKEGSGFGLTGRGDSFVLIAAVIPGSAAGLKEGDIYIVSNGQPCRW
RHAEEVTLKAGASLQVSLSPSSRLGDRPVLILGPRLLRSOREHGCKTP
ASTWASPRLLNRSRKAQQKGTGGCPQCPAPKVPAPPPSLKHPGP"

misc_feature

204. .341

/note="HRI; Region: Hri repeat"

/db_xref="CDD:pfam02185"

423. .881

misc_feature

/note="BRO1; Region: BRO1-like domain. This functionally uncharacterised domain is found in a number of signal

transduction proteins, including Rophalin and BRO1"

/db_xref="CDD:pfam03097"

misc_feature

1641. .1841

/note="PDZ; Region: Domain present in PSD-95, Dlg, and ZO-1/2. Also called DHR (Dlg homologous region) or GIGF (relatively well conserved tetrapeptide in these domains). Some PDZs have been shown to bind C-terminal polypeptides"

/db_xref="CDD:smart00228"

ORIGIN

Query Match 85.3%; Score 2105; DB 9; Length 3763;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2116; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 321 GGGCTGTGACTCCCTGAGCGCAGATCCAGTGGCGCCAGCTCAGAGCGCGAGGCCAGAT 380
Db 161 GGGCTGTGACTCCCTGAGCGCAGATCCAGTGGCGCCAGCTCAGAGCGCGAGGCCAGAT 220
Qy 381 TCACCAGCAGATTGACAAAGGAGCTGCAGATGCGGAGCGGCGCTGAGAACCTCTACAGAGC 440
Db 221 TCACCAGCAGATTGACAAAGGAGCTGCAGATGCGGAGCGGCGCTGAGAACCTCTACAGAGC 280
Qy 441 CACAGACAAACCGGGTGAAGAGAGCGTGCCTGGAGCTGAGCTAGCTACGTCACTCAA 500
Db 281 CACCAGCAACAACCGGGTGAAGAGAGCGTGCCTGGAGCTGAGCTAGCTCACTCAA 340
Qy 501 CCTGCACTGCTCAAGAGGAGGAGCTGGAGGAGCTCAGCGTGGGTGGACCTGGCGCGCA 560
Db 341 CCTGCACTGCTCAAGAGGAGGAGCTGGAGGAGCTCAGCGTGGGTGGACCTGGCGCGCA 400
Qy 561 TGGGAGCGAAGCTGTCACTGTCCCATGATCCCTGGGCTGAAGGAGACCAAGGAGCT 620
Db 401 TGGGAGCGAAGCTGTCACTGTCCCATGATCCCTGGGCTGAAGGAGACCAAGGAGCT 460
Qy 621 G3ACTGTCTACACCGCTGAAGAGCTGATCTCAGTGCACCTTTGGAGAGAGCGCGCTC 680
Db 461 G3ACTGTCTACACCGCTGAAGAGCTGATCTCAGTGCACCTTTGGAGAGAGCGCGCTC 520
Qy 681 CTRAGGAGCAGAAATCAGGAGCTGGAGGCGCTTGGCGAGGCCATGCGGACCCCGACCG 740
Db 521 CTRAGGAGCAGAAATCAGGAGCTGGAGGCGCTTGGCGAGGCCATGCGGACCCCGACCG 580
Qy 741 GAATGATCGGGCTGGAGCTGCTCAAGCCTATTACAAACAGCTGTCTTCTTGGATGC 800
Db 581 GAATGATCGGGCTGGAGCTGCTCAAGCCTATTACAAACAGCTGTCTTCTTGGATGC 640
Qy 801 GCGCTTCTCACCGCTGCAGAGGCTCGGGCTCTTCTCCACTGGTACGACTCGCTTAC 860
Db 641 GCGCTTCTCACCGCTGCAGAGGCTCGGGCTCTTCTCCACTGGTACGACTCGCTTAC 700
Qy 861 TGGGGTCCCGGCGCAGCAGCTGCGCTTGGCGCTTCGAGAAGCGGAGCGCTTCTTCAACAT 920
Db 701 TGGGGTCCCGGCGCAGCAGCTGCGCTTGGCGCTTCGAGAAGCGGAGCGCTTCTTCAACAT 760
Qy 921 CGGTGCGCTTCCACAACGAGATTGGGGCGCGCAGGACCGCTCTCTGCAACGAGGTGCGCG 980
Db 761 CGGTGCGCTTCCACAACGAGATTGGGGCGCGCAGGACCGCTCTCTGCAACGAGGTGCGCG 820
Qy 981 CCGGGCTATCGAGGCGCTTCCAGAGGCGCTGGGGCTTCCAGCTCTCTGAGGAGGAACTT 1040
Db 821 CCGGGCTATCGAGGCGCTTCCAGAGGCGCTGGGGCTTCCAGCTCTCTGAGGAGGAACTT 880
Qy 1041 CTCCTATGCGCGAGCCAGACATGAGCGCTGGCTCCCTCTGCGCACTGGAGCAGCTCAT 1100
Db 881 CTCCTATGCGCGAGCCAGACATGAGCGCTGGCTCCCTCTGCGCACTGGAGCAGCTCAT 940
Qy 1101 GATGGCCAGCGCCAGGAATGTGTGTTGAGGCGCTCTCAACCTCTCCCTCATGGCGCC 1160
Db 941 GATGGCCAGCGCCAGGAATGTGTGTTGAGGCGCTCTCAACCTCTCCCTCATGGCGCC 1000
Qy 1161 CCAAGACTGCTGCGCCAGCTGGCGCTGGCGAGGCGCCCGAGCTGGAGCGCGAGTA 1220
Db 1001 CCAAGACTGCTGCGCCAGCTGGCGCTGGCGAGGCGCCCGAGCTGGAGCGCGAGTA 1060

1221 CAGGCTAGTGCACCGAGACCATGCGCCAGCACCCGCTCCACGACTAGTGCCTGTCTCCTG 1280
1061 CAGGCTAGTGCACCGAGACCATGCGCCAGCACCCGCTCCACGACTAGTGCCTGTCTCCTG 1120
1281 GACTGCCCTGCTGATGTCAGGCGGAGTACTTCGCTCCCTGCGCCACTACACGATAGC 1340
1121 GACTGCCCTGCTGATGTCAGGCGGAGTACTTCGCTCCCTGCGCCACTACACGATAGC 1180
1341 CATGCCCTCTGCGAGCGCTCCCGAGCGACCGAGGAGAGCTCCCGACCGACGAGCAGGT 1400
1181 CATGCCCTCTGCGAGCGCTCCCGAGCGACCGAGGAGAGCTCCCGACCGACGAGCAGGT 1240
1401 CTTCTGACGCCCCCCTCCTCTTAAGCCCCGAGGCCCTGTGCTGCCGACGAGCTGGA 1460
1241 CTTCTGACGCCCCCCTCCTCTTAAGCCCCGAGGCCCTGTGCTGCCGACGAGCTGGA 1300
1461 GGAGCGCAGGAGCTGGCAAGGCACCTGAAGCTGTCATCTCTGGGCGCAGGAGGCG 1520
1301 GGAGCGCAGGAGCTGGCAAGGCACCTGAAGCTGTCATCTCTGGGCGCAGGAGGCG 1360
1521 GCTGCGGCTGACGCGCTCTGCGCGCTCTGCGCGAGGTGACCTCTTCCGGCTGTGAT 1580
1361 GCTGCGGCTGACGCGCTCTGCGCGCTCTGCGCGAGGTGACCTCTTCCGGCTGTGAT 1420
1581 CTTCCAGAGCTGACGCGCTCACTGCGCAAGTATGCGGAGCTCGACCGTGAGGATGACTT 1640
1421 CTTCCAGAGCTGACGCGCTCACTGCGCAAGTATGCGGAGCTCGACCGTGAGGATGACTT 1480
1641 CTGTGAGCTGCGGAGGCCCGGACATCCAGCTTAAGACCCAGACCCAGAGGCGGAG 1700
1481 CTGTGAGCTGCGGAGGCCCGGACATCCAGCTTAAGACCCAGACCCAGAGGCGGAG 1540
1701 GATGCAAGCTGCTCCAGGGAAGGGCTGACATCTTCCATCGCTGGGCGCCCTGTC 1760
1541 GATGCAAGCTGCTCCAGGGAAGGGCTGACATCTTCCATCGCTGGGCGCCCTGTC 1600
1761 TGTGTTCTCAGCAAGACCGTGGCGGTGTGTGGGCGCCGTCACCTGACCCGAGGAGA 1820
1601 TGTGTTCTCAGCAAGACCGTGGCGGTGTGTGGGCGCCGTCACCTGACCCGAGGAGA 1660
1821 GGGCGGCTTGGCTCAGCTTGGGGGAGCTCGCTGCTCATCGCTGCGCTCATTC 1880
1661 GGGCGGCTTGGCTCAGCTTGGGGGAGCTCGCTGCTCATCGCTGCGCTCATTC 1720
1881 AGGGAGCCAGGCGCGGCGGCTGGCTGAAGGAGGCGCACTACATTGTGTCAAGTGAATGG 1940
1721 AGGGAGCCAGGCGCGGCGGCTGGCTGAAGGAGGCGCACTACATTGTGTCAAGTGAATGG 1780
1941 GCAGCCATGCAAGTGTGAGACACGCGAGGTGTGACGAGCTGAAGGCTGCGGGAGA 2000
1781 GCAGCCATGCAAGTGTGAGACACGCGAGGTGTGACGAGCTGAAGGCTGCGGGAGA 1840
2001 GCGGGCGCCAGCTGCAAGTGTGCTGCTGCGGAGCTGACCTGACACTGCGCCAGCTTGGG 2060
1841 GCGGGCGCCAGCTGCAAGTGTGCTGCTGCGGAGCTGACCTGACACTGCGCCAGCTTGGG 1900
2061 GCGGGCGCCAGCTGCAAGTGTGCTGCTGCGGAGCTGACCTGACACTGCGCCAGCTTGGG 2120
1901 GCGGGCGCCAGCTGCAAGTGTGCTGCTGCGGAGCTGACCTGACACTGCGCCAGCTTGGG 1960
2121 TTGCAAGACCCCGGATCCAGTGGGCGAGTCCCGGCGCCCTCTCAACTGAGCGCGAAA 2180
1961 TTGCAAGACCCCGGATCCAGTGGGCGAGTCCCGGCGCCCTCTCAACTGAGCGCGAAA 2020
2181 GCGCCAGCAGGCGCAGACTGAGGCTGCGCCAGCGCTGTGCGCCAGCTGAAGCCAGCTCC 2240
2021 GCGCCAGCAGGCGCAGACTGAGGCTGCGCCAGCGCTGTGCGCCAGCTGAAGCCAGCTCC 2080
2241 GCGCTCATCTTGAAGCACCCAGGCTGGCGCTGAGGGCCAGGATCCCTGCAGCGCCCTCAG 2300
2081 GCGCTCATCTTGAAGCACCCAGGCTGGCGCTGAGGGCCAGGATCCCTGCAGC-CCTCAG 2139

2301 CCTGGCTCCAGTGGCAGCAAGCAGCAGCATGCCCTCCACCAGAGGACCTCCGGG 2360
2140 CCTGGCTCCAGTGGCAGCAAGCAGCAGCATGCCCTCCACCAGAGGACCTCCGGG 2199
2361 CAATGCTGTGCTGGCTCATGCTGAGGCTGCTGGGACCTGCTGCCCATTAAGAC 2420
2200 CAATGCTGTGCTGGCTCATGCTGAGGCTGCTGGGACCTGCTGCCCATTAAGAC 2259
2421 TGGTCAGACCTGTCTGA 2437
2260 TGGTCAGACCTGTCTGA 2276

RESULT 5
AB067516
LOCUS AB067516 3647 bp mRNA linear PRI 15-SEP-2001
DEFINITION Homo sapiens mRNA for KIAA1929 protein, partial cds.
ACCESSION AB067516
VERSION AB067516.1 GI:15620916
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Nagase, T., Kikuno, R. and Ohara, O.
Prediction of the coding sequences of unidentified human genes.
XXI. The complete sequences of 60 new cDNA clones from brain which
code for large proteins
DNA Res. 8 (4), 179-187 (2001)
21456161
11572484

REFERENCE
JOURNAL
MEDLINE
PUBMED
REFERENCE
2 (bases 1 to 3647)
AUTHORS
Ohara, O., Nagase, T. and Kikuno, R.
Direct Submission
TITLE
JOURNAL
Submitted (27-JUL-2001) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
URL: http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,
Fax:81-438-52-3914)

FEATURES
Location/Qualifiers
1..3647
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="fj08985"
/tissue_type="brain"
/note="vector:pBluescriptII SK plus"
1..3647
/gene="KIAA1929"
<1..1233
/gene="KIAA1929"
/note="Start codon is not identified. fj08985 cDNA clone
for KIAA1929 has 349-bp, 320-bp and 413-bp insertions
after the positions 450, 936 and 1283 of the sequence of
KIAA1929, respectively."
/codon_start=1
/product="KIAA1929 protein"
/protein_id="BAB67822.1"
/db_xref="GI:15620917"

gene
CDS
/translation="EERPVGAGAGEEPRLGQCDLSLTQIQGQLQSRRAQTHQIDKE
LQMTGAENLYRATNNRIVALELSYVNSLQLKELELSELSGGVDPGRHSEAV
TVPMIPLGKETKELDWLPLKLSVHFVFDGASVBAEIRLEALRAQMTPSRNE
GLELTYNOLCFDARFLTPARSLGLFPHWYDLSLTGPAQALAPKSGVLPNIG
ALHTQIGARQDRSCTEGARRAMEAPQRAAGAFSLIRENFSPDPDMSAASICAL
EQ
MMAQECFEGLSFPFASMAPQCLQLRLAQEAQVAAYRLVHRTNAQPPVHDYVP
VSWALVHVKAAYFRSLAHYHVMALCDGSRCECPHLPVLPVLPVLPVLPVLPVLPV
RGRPHARAGLPAAPHLL"

ORIGIN
Query Match 81.8%; Score 2019; DB 9; Length 3647;
Best local Similarity 96.5%; Pred. No. 0;
Matches 2116; Conservative 0; Mismatches 0; Indels 77; Gaps 2;

Qy	321	GGGCTGTGACTCCCTGACGCGAGATCCAGTGCGGCCAGCTGCAGAGCGCGAGGGCCAGAT	380	Qy	1364	-----CAGCGACCGAGGAGAGCTCC	1384
Db	51	GGGCTGTGACTCCCTGACGCGAGATCCAGTGCGGCCAGCTGCAGAGCGCGAGGGCCAGAT	110	Db	1131	GGCCCCCGCGCAGGGCTCACAGCCTCTGTGTCCCCAGCAGCGAGCGAGAGAGCTCC	1190
Qy	381	TCACCAAGCATTGACAAGGAGCTGCAGATGCGGACGGGGCGCTGAGAACTCTTACAGAGC	440	Qy	1385	CCACGACGACGAGGTTTCTCTCAGCCCCCACTCTCTAAAGCCCCGAGGGCCCTGTGC	1444
Db	111	TCACCAAGCATTGACAAGGAGCTGCAGATGCGGACGGGGCGCTGAGAACTCTTACAGAGC	170	Db	1191	CCACGACGACGAGGTTTCTCTCAGCCCCCACTCTCTAAAGCCCCGAGGGCCCTGTGC	1250
Qy	441	CACACGAAACAAACGGGTGAGAGACGGTGCGCCCTGGAGCTGAGCTGACTCAACTCAA	500	Qy	1445	TGCGCAGBAGCTGAGGAGCGCAGCAGCTTGCGAAGGCACACCTGAGAGGTGCCATCC	1504
Db	171	CACACGAAACAAACGGGTGAGAGACGGTGCGCCCTGGAGCTGAGCTGACTCAACTCAA	230	Db	1251	TGCGCAGBAGCTGAGGAGCGCAGCAGCTTGCGAAGGCACACCTGAGAGGTGCCATCC	1310
Qy	501	CTTGCGAGCTGCTGAAGGAGAGCTGGAGGAGCTCAGCGGTGGCTGACACCTGGCGGCA	560	Qy	1505	TGCGGAGGAGGAGGCGCTGCGGCTGCAAGCTGTCGCCGCTCTCTGCGGAGGTGAGCC	1564
Db	231	CTTGCGAGCTGCTGAAGGAGAGCTGGAGGAGCTCAGCGGTGGCTGACACCTGGCGGCA	290	Db	1311	TGCGGAGGAGGAGGCGCTGCGGCTGCAAGCTGTCGCCGCTCTCTGCGGAGGTGAGCC	1370
Qy	561	TGGGAGCGAAGCTGTCACTGTCCCATGATCCGCCCTGGGCTGAAGGAGACCAAGAGACT	620	Qy	1565	TGCTTGGGCTGTGATCTCCACAGCTGACGCGCTCACTGGCCAACTATGCGGAGGTGCG	1624
Db	291	TGGGAGCGAAGCTGTCACTGTCCCATGATCCGCCCTGGGCTGAAGGAGACCAAGAGACT	350	Db	1371	TGCTTGGGCTGTGATCTCCACAGCTGACGCGCTCACTGGCCAACTATGCGGAGGTGCG	1430
Qy	621	GGACTGGTCTACACCGCTGAAGGAGCTGATCTCAGTGCACTTTGGAGAGGACGGCGCTC	680	Qy	1625	ACCGTGAGGATGACTTCTGTGAGGCTGCCGAGGCCCCGGACATCCAGCCTAAGACCCACC	1684
Db	351	GGACTGGTCTACACCGCTGAAGGAGCTGATCTCAGTGCACTTTGGAGAGGACGGCGCTC	410	Db	1431	ACCGTGAGGATGACTTCTGTGAGGCTGCCGAGGCCCCGGACATCCAGCCTAAGACCCACC	1490
Qy	681	CTACGAGGCAAGAAATCAGGAGCTGGAGGCGCTGCCGAGGSCCATGCGGACCCCGCCG	740	Qy	1685	AGAAGCCACAGGCGCAGGATGCCACGCGCTGTCCACAGGGGAAAGGGGCCCTGACATCTTCCATC	1744
Db	411	CTACGAGGCAAGAAATCAGGAGCTGGAGGCGCTGCCGAGGSCCATGCGGACCCCGCCG	470	Db	1491	AGAAGCCACAGGCGCAGGATGCCACGCGCTGTCCACAGGGGAAAGGGGCCCTGACATCTTCCATC	1550
Qy	741	GAATGAGTGGGGCTGGAGCTGTCTCAGAGCTATTACACAGCTGTGTCTCTGATGC	800	Qy	1745	GGCTGGGGCCCCCTGTCTGTGTTCTCAGCAGAAACCGGTGGCGGCTGGTGGGGCCCGTCC	1804
Db	471	GAATGAGTGGGGCTGGAGCTGTCTCAGAGCTATTACACAGCTGTGTCTCTGATGC	530	Db	1551	GGCTGGGGCCCCCTGTCTGTGTTCTCAGCAGAAACCGGTGGCGGCTGGTGGGGCCCGTCC	1610
Qy	801	GGGCTTCTCAGCCCTGCGAGAGCTGCGGGCTCTTCTTCACAGTGCTGAGCTGCTTAC	860	Qy	1805	ACCTGACCCGAGAGAGGGCGGCTTTGGCCTCAGCTTCGSGGAGACTGCGCTGTCTCTCA	1864
Db	531	GGGCTTCTCAGCCCTGCGAGAGCTGCGGGCTCTTCTTCACAGTGCTGAGCTGCTTAC	590	Db	1611	ACCTGACCCGAGAGAGGGCGGCTTTGGCCTCAGCTTCGSGGAGACTGCGCTGTCTCTCA	1670
Qy	861	TGGGGTCCCGGCCAGCAGCTGCCCTGGCTTCGAGAGGGAGCGCTCTCTTCAACAT	920	Qy	1865	TCGCTGCGCTTCATTCACGAGGAGCAGGCGCGGCGCTGGCCTCTGAAGAGGGGCGACTACA	1924
Db	591	TGGGGTCCCGGCCAGCAGCTGCCCTGGCTTCGAGAGGGAGCGCTCTCTTCAACAT	650	Db	1671	TCGCTGCGCTTCATTCACGAGGAGCAGGCGCGGCGCTGGCCTCTGAAGAGGGGCGACTACA	1730
Qy	921	CGGTGCCCTCCACACGAGATTGGGGCGCGCCAGGACCGCTCTGCAACGAGGAGTGC	980	Qy	1925	TTGTGTCAGTGAATGGCAGCCATGCAAGTGTGTGAGACACGCGGAGGTGCTGACGAGC	1984
Db	651	CGGTGCCCTCCACACGAGATTGGGGCGCGCCAGGACCGCTCTGCAACGAGGAGTGC	710	Db	1731	TTGTGTCAGTGAATGGCAGCCATGCAAGTGTGTGAGACACGCGGAGGTGCTGACGAGC	1790
Qy	981	CCGGCTATCGAGGCTTTCAGAGGCGCTGGGGCTTCAGCCTCTGAGGAGAACTT	1040	Qy	1985	TGAAGGCTGGGAGAGGGCGGCGCAGCTGTCAGAGTGTGTGCTGCTGCCAGCTCTA	2044
Db	711	CCGGCTATCGAGGCTTTCAGAGGCGCTGGGGCTTCAGCCTCTGAGGAGAACTT	770	Db	1791	TGAAGGCTGGGAGAGGGCGGCGCAGCTTCAAGTGTGTGCTGCTGCCAGCTCTA	1850
Qy	1041	CTCCATGCGCGAGCCAGACATGAGCGCTGCGTCCCTCTGCGCATGAGAGCAGCTCAT	1100	Qy	2045	GACTGCCCAAGCTTTGGGGACCGCGCGCCCGTCTGTGTGGCCCCCAGGGGGCTTCTAAGGA	2104
Db	771	CTCCATGCGCGAGCCAGACATGAGCGCTGCGTCCCTCTGCGCATGAGAGCAGCTCAT	830	Db	1851	GACTGCCCAAGCTTTGGGGACCGCGCGCCCGTCTGTGTGGCCCCCAGGGGGCTTCTAAGGA	1910
Qy	1101	GATGGCCAGGCCAGGAATGTGTGTTGAGGGCTCTCAACAGCTGCTCCATGGCCCC	1160	Qy	2105	GCACAGGGAGCATGTTGCAAGACCCCGGCATCCAGTGGGCGCAGTCCCGGGCCCCCTCC	2164
Db	831	GATGGCCAGGCCAGGAATGTGTGTTGAGGGCTCTCAACAGCTGCTCCATGGCCCC	890	Db	1911	GCACAGGGAGCATGTTGCAAGACCCCGGCATCCAGTGGGCGCAGTCCCGGGCCCCCTCC	1970
Qy	1161	CCAGACTGCTGGCCAGCTGCGCTGGCGCAGGAGCGCGCCAGGTGGCAGCCGAGTA	1220	Qy	2165	TCAACTGGAGCCCAAAGGCCCAGCAGGGCAAGACTGGAGGCTGCCCCAGCCTGTGCC	2224
Db	891	CCAGACTGCTGGCCAGCTGCGCTGGCGCAGGAGCGCGCCAGGTGGCAGCCGAGTA	950	Db	1971	TCAACTGGAGCCCAAAGGCCCAGCAGGGCAAGACTGGAGGCTGCCCCAGCCTGTGCC	2030
Qy	1221	CAGGCTAGTGACCGGACCAATGGGCCAGCCAGCCGCTCCAGACTACGTGCTGCTCTG	1280	Qy	2225	CAGTGAAGCAGCTTCGCGCTCATCTTTGAAGCACCCAGGGTGGCGCTGAGGGCAGGAT	2284
Db	951	CAGGCTAGTGACCGGACCAATGGGCCAGCCAGCCGCTCCAGACTACGTGCTGCTCTG	1010	Db	2031	CAGTGAAGCAGCTTCGCGCTCATCTTTGAAGCACCCAGGGTGGCGCTGAGGGCAGGAT	2090
Qy	1281	GACTGCCCTGGTGAATGCAAGGCGGAGTACTTCCGCTCCCTGGCCCACTACACGATAGC	1340	Qy	2285	CCCTGACGCGCTCAGCGCTGGCTCCAGCTGGCAGCAAGCACCCAGCATGCCCTCCCCAC	2344
Db	1011	GACTGCCCTGGTGAATGCAAGGCGGAGTACTTCCGCTCCCTGGCCCACTACACGATAGC	1070	Db	2091	CCCTGACG--CTCAGCGCTGGCTCCAGCTGGCAGCAAGCACCCAGGAGTGGCCCTCCCCAC	2149
Qy	1341	CATGGCCCTCTGCGACGCTCCC-----	1363	Qy	2345	CCAGAGGACCTCCGGGCAATGCTGCTCCGCTCATGCTGAGGCTGCTCTGGGCACTG	2404
Db	1071	CATGGCCCTCTGCGACGCTCCC-----	1130	Db	2150	CCAGAGGACCTCCGGGCAATGCTGCTCCGCTCATGCTGAGGCTGCTCTGGGCACTG	2209
				Qy	2405	CCTGCCCAATTAAAGACTGGTTCAGACCTGTCTGA	2437

```
Db      2210 CCTGCCATTAAAGACTGGTACAGCTGTCTGA 2242
|||||
RESULT 6
LOCUS   AY082588      2088 bp mRNA linear PRI 12-NOV-2002
DEFINITION Homo sapiens rhophilin-1 mRNA, complete cds.
ACCESSION AY082588
VERSION   AY082588.1 GI:19697912
KEYWORDS
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2088)
AUTHORS   Peck, J.W., Oberst, M., Bouker, K.B., Bowden, E. and Burbelo, P.D.
TITLE     The RhoA-binding protein, Rhophilin-2, Regulates Actin Cytoskeleton
          Organization
JOURNAL   J. Biol. Chem. 277 (46), 43924-43932 (2002)
PUBMED   1221077
REFERENCE 2 (bases 1 to 2088)
AUTHORS   Burbelo, P.D.
TITLE     Structure and function of Rhophilin homologs
JOURNAL   Unpublished
REFERENCE 3 (bases 1 to 2088)
AUTHORS   Burbelo, P.D.
TITLE     Direct Submission
JOURNAL   Submitted (06-MAR-2002) Dept. of Oncology, Georgetown University
          Medical Center, 3970 Reservoir Rd., N.W., Washington, DC 20007, USA
FEATURES
         Location/Qualifiers
         1..2088
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:2155941"
         1..2088
            /note="non-kinase effector protein of RhoA; RhoA-binding
            protein"
         /codon_start=1
         /product="rhophilin-1"
         /protein_id="AAI89809.1"
         /db_xref="GI:19697913"
         /translation="MILREPDGAGEEPRLOGDSLTQIQQLQSRRAQIHQOI
         DKELQMTGAEINRYATSNRRVETVALELSYVNSNLQILKEELEELSGVDYPSRHS
         EAVTVMPLGLKETKEDLDWSTPLKELISVHFEDGSAIELEALROAMRTPSR
         NESGLELTAYINOLCFDARFLTPARSLGLFFHWYDSLTGVPQORALAFKGSVLFP
         NIGALHTOIGARQDRSCTEGARAMEAFORAAAGAFSLURENFHAPSPDMSASICAL
         BOLMAQAEQCVFSGLSPPASAPODCLAQLRLAQRAQVAAYELVHRITWQPPVHD
         YVPVSWTALVHVKAIEFRSLAHYHVMALDGSRECPHLMVJPRPRAGSPQLCPP
         AATGSELTHQVFLQPTSPKPRGPVLVQLEERROIGKAHLKRALILQBEALRLHA
         LCVRLREVLLURAVISQTLQSLAKYAELEDREDDFCEAAEAPDIQPKTHQPEARMPR
         LSRQKGEVDIFHRLGLPLSVFSANRWLVGPHVLTFRGGFGLTLRGDSPLVIAAVIPG
         SQAAAAGLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAGEAGASLQVWSLLPSRLPSL
         GDRPVLILGPRGLRSQREHGCKTPASTWASPRPLNWSRAQQGKTGGCQPCAPVK
         PAPSSLLKHPWP"
ORIGIN
Query Match 75.7%; Score 1869; DB 9; Length 2088;
Best Local Similarity 96.3%; Pred. No. 5,4e-298;
Matches 1954; Conservative 0; Mismatches 0; Indels 75; Gaps 1;
QY      321 GGGCTGTGACTCCCTGAGCGAGATCCAGTGGCGGCAGCTGCAGAGCCGAGGGCCAGAT 380
Db      60 GGGCTGTGACTCCCTGAGCGAGATCCAGTGGCGGCAGCTGCAGAGCCGAGGGCCAGAT 119
QY      381 TCACAGAGATTGACAGGAGCTGCAGATGGGAGCGGCGCTGAGAACCTTACAGAGC 440
Db      120 TCACAGAGATTGACAGGAGCTGCAGATGGGAGCGGCGCTGAGAACCTTACAGAGC 179
QY      441 CACCAGCAACACCGGGTGGAGAGAGCGGTGCGCCCTGGAGCTGAGCTACGTCAACTCCAA 500
Db      180 CACCAGCAACACCGGGTGGAGAGAGCGGTGCGCCCTGGAGCTGAGCTACGTCAACTCCAA 239
QY      501 CCTGCAGCTGCTGAAGAGGAGAGCTGAGGAGAGCTCAGCGGTGGCGTGGACCTGGCGCGCA 560
Db      240 CCTGCAGCTGCTGAAGAGGAGAGCTGAGGAGAGCTCAGCGGTGGCGTGGACCTGGCGCGCA 299
QY      561 TGGGAGCGGAAGCTGTCACTGTCCCATGATGATCCCTTGGGCTTGAAGAGACCAAGAGCT 620
Db      300 TGGGAGCGGAAGCTGTCACTGTCCCATGATGATCCCTTGGGCTTGAAGAGACCAAGAGCT 359
QY      621 GGACTGCTCTACACCGCTGAAGGAGCTGATCTCAGTGCACATTTGGAGAGAGCGGCGCTC 680
Db      360 GGACTGCTCTACACCGCTGAAGGAGCTGATCTCAGTGCACATTTGGAGAGAGCGGCGCTC 419
QY      681 CTACGAGGAGAGAAATCAGGAGCTGAGGAGCTCGGCGAGGCGCATGCGGACCCCGAGCGG 740
Db      420 CTACGAGGAGAGAAATCAGGAGCTGAGGAGCTCGGCGAGGCGCATGCGGACCCCGAGCGG 479
QY      741 GAATGAGTCGGGCTTGGAGCTGCTCAGAGCTATTATCAACACAGCTGTGCTTCTGGATGC 800
Db      480 GAATGAGTCGGGCTTGGAGCTGCTCAGAGCTATTATCAACACAGCTGTGCTTCTGGATGC 539
QY      801 GCGCTTCTCTACCCCTGCGAGGAGCTCGGGCTCTTCTTCCACTGCTAGCTCGCTTAC 860
Db      540 GCGCTTCTCTACCCCTGCGAGGAGCTCGGGCTCTTCTTCCACTGCTAGCTCGCTTAC 599
QY      861 TGGGCTCCCGGCCAGCAGCGTGGCTTCCGAGAGAGGCGAGCGTTCCTTCAACAT 920
Db      600 TGGGCTCCCGGCCAGCAGCGTGGCTTCCGAGAGAGGCGAGCGTTCCTTCAACAT 659
QY      921 CGGTGCTCTCAGACGAGATTGGGGCGGCGCAGACCGCTCTCGACCGAGGGTGCCTCG 980
Db      660 CGGTGCTCTCAGACGAGATTGGGGCGGCGCAGACCGCTCTCGACCGAGGGTGCCTCG 719
QY      981 CCGGCTATGAGAGGCTTCCAGAGGCGCGCTGGGGCTTCCAGCTTCTGAGGAGAACTT 1040
Db      720 CCGGCTATGAGAGGCTTCCAGAGGCGCGCTGGGGCTTCCAGCTTCTGAGGAGAACTT 779
QY      1041 CTCCATGCGCGGCCAGACATGAGCGTGGCTTCTGCGCATCTGGAGCAGCTCAT 1100
Db      780 CTCCATGCGCGGCCAGACATGAGCGTGGCTTCTGCGCATCTGGAGCAGCTCAT 839
QY      1101 GATGCGCCAGGCGCCAGGAATGTGTTTGGAGGCTCTCAACACCTGCTCCATGGCCCC 1160
Db      840 GATGCGCCAGGCGCCAGGAATGTGTTTGGAGGCTCTCAACACCTGCTCCATGGCCCC 899
QY      1161 CCAAGACTGCTGCGCCAGCTGCGCTGCGCGCAGAGAGCGCCCGCAGGTGGCAGCGAGTA 1220
Db      900 CCAAGACTGCTGCGCCAGCTGCGCTGCGCGCAGAGAGCGCCCGCAGGTGGCAGCGAGTA 959
QY      1221 CAGGCTAGTGACCGGACCATGGCCAGCCACCGCTCCACGACTACGTGCTGTCTCTCTG 1280
Db      960 CAGGCTAGTGACCGGACCATGGCCAGCCACCGCTCCACGACTACGTGCTGTCTCTCTG 1019
QY      1281 GACTGCTGTGTGATGTCAAGGCGGAGTACTTGGCTTCTGCGCCCTACCTACCGAGTAC 1340
Db      1020 GACTGCTGTGTGATGTCAAGGCGGAGTACTTGGCTTCTGCGCCCTACCTACCGAGTAC 1079
QY      1341 CATGCGCCCTCTGCGAGGCTCTCC----- 1363
Db      1080 CATGCGCCCTCTGCGAGCGCTCTCCGTGAGTGGCCACCAACATTTGGCCATGTGTACG 1139
QY      1364 ----- 1385
Db      1140 GCCCCCCCGCGCAGGCTCACAGCCTCTCTGTGTCCTCCAGCAGCAGCGAGGAGAGCTCCC 1199
QY      1386 CACGCAAGAGAGAGTCTTCTGTGAGCCCCCCTCTCTCTTAAGCCCCGAGGCGCTGTGCT 1445
Db      1200 CACGCAAGAGAGTCTTCTGTGAGCCCCCCTCTCTCTTAAGCCCCGAGGCGCTGTGCT 1259
QY      1446 GCCGAGAGCTGGAGAGGCGCAGGAGCTTGGCAAGGCGCACCTGAGCGGTGCGCATCT 1505
Db      1260 GCCGAGAGCTGGAGAGGCGCAGGAGCTTGGCAAGGCGCACCTGAGCGGTGCGCATCT 1319
```

QY 1506 GGGCAGGAGGAGCGCTGGCTGCACGCCCTGTCGCGGTCTCGCGAGGTGACCT 1565
DB 1320 GGGCAGGAGGAGCGCTGGCTGCACGCCCTGTCGCGGTCTCGCGAGGTGACCT 1379
QY 1566 GCTTCGGGCTGTGATCTCCAGACGCTGACGCGCTCACTGGCCAAAGTATCGGAGGTGCA 1625
DB 1380 GCTTCGGGCTGTGATCTCCAGACGCTGACGCGCTCACTGGCCAAAGTATCGGAGGTGCA 1439
QY 1626 CCGTGAGGATGATCTCTGTGAGGCTGCGAGGCGCCGGAATCAGCGCTTAAGACCCACCA 1685
DB 1440 CCGTGAGGATGATCTCTGTGAGGCTGCGAGGCGCCGGAATCAGCGCTTAAGACCCACCA 1499
QY 1686 GAGCCAGAGGCGAGGATGACCGCTGTCAGAGGAGGAGGCGCTGACATCTTCATCG 1745
DB 1500 GAGCCAGAGGCGAGGATGACCGCTGTCAGAGGAGGAGGCGCTGACATCTTCATCG 1559
QY 1746 GCTGGGCGCCCTGTCTGTGTTCTCAGCAAGAACCGGTGGCGGTGGTGGGCGCCGTCGA 1805
DB 1560 GCTGGGCGCCCTGTCTGTGTTCTCAGCAAGAACCGGTGGCGGTGGTGGGCGCCGTCGA 1619
QY 1806 CTTGACCCGAGGAGGCGCGCTTTGGCTCAGCTTCGGGGAGACTCGCTGTCTCAT 1865
DB 1620 CTTGACCCGAGGAGGCGCGCTTTGGCTCAGCTTCGGGGAGACTCGCTGTCTCAT 1679
QY 1866 CGCTGCCGTCTATCCAGGGAGCCAGGCGCGCGGCTGGCTGGAAGGAGGCGCACTACAT 1925
DB 1680 CGCTGCCGTCTATCCAGGGAGCCAGGCGCGCGGCTGGCTGGAAGGAGGCGCACTACAT 1739
QY 1926 TGTGTGATGAATGGGAGCCATGACAGGTGGTGGAGACACCGCGAGGTGGTGGAGGCT 1985
DB 1740 TGTGTGATGAATGGGAGCCATGACAGGTGGTGGAGACACCGCGAGGTGGTGGAGGCT 1799
QY 1986 GAGGCTGCGGAGAGGCGCGCGCTGACAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2045
DB 1800 GAGGCTGCGGAGAGGCGCGCGCTGACAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1859
QY 2046 ACTGCCAGCTTGGGGAGCGCGCGCTGCTGCTGGCGCCAGGGGGCTTCTAAGGAG 2105
DB 1860 ACTGCCAGCTTGGGGAGCGCGCGCTGCTGCTGGCGCCAGGGGGCTTCTAAGGAG 1919
QY 2106 CCAGAGGAGCATGGTTCGAAGACCCCGGATCCAGCTGGGCGCAAGTCCCGGCGCCCTCT 2165
DB 1920 CCAGAGGAGCATGGTTCGAAGACCCCGGATCCAGCTGGGCGCAAGTCCCGGCGCCCTCT 1979
QY 2166 CAATGAGGCGAAGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2225
DB 1980 CAATGAGGCGAAGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2039
QY 2226 AGTGAAGCCAGCTCGCGCTCATCTTGAAGCACCCAGGGTGGCGGTGA 2274
DB 2040 AGTGAAGCCAGCTCGCGCTCATCTTGAAGCACCCAGGGTGGCGGTGA 2088

RESULT 7
MMU43194 1932 bp mRNA linear ROD 02-FEB-1996
LOCUS Mus musculus rhophilin mRNA, complete cds.
DEFINITION U43194
ACCESSION U43194.1 GI:1176421
VERSION rho binding protein; PKN homology.
KEYWORDS Mus musculus (house mouse)
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1932)
Watanabe,G., Saito,Y., Madaule,P., Ishizaki,T., Fujisawa,K.,
Mori,N., Mukai,H., Ono,Y., Kakizuka,A. and Narumiya,S.
TITLE Protein kinase N (PKN) and PKN-related protein rhophilin as targets
of small GTPase Rho
JOURNAL Science 271 (5249), 645-648 (1996)
MEDLINE 96165390
PUBMED 8571126
REFERENCE 2 (bases 1 to 1932)

AUTHORS Watanabe,G. and Narumiya,S.
DIRECT SUBMISSION
JOURNAL TITLE Direct Submission
JOURNAL Submitted (13-DEC-1995) Go Watanabe, Pharmacology, Kyoto University
of Fac. Med., Yoshida Konoe-Cho Sakyo-Ku, Kyoto City, 606, Japan
FEATURES
Location/Qualifiers
1..1932
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..1932
/function="small GTPase Rho binding protein"
/codon_start=1
/product="rhophilin"
/protein_id="AAC52388.1"
/db_xref="GI:1176422"
/translation="MILERPQCGTGESSRRPDDGSIKRGYGVQNPQGOLOSHR
ARLHQIGIKELRMRTAEMLYRATSNWVETVALSIVNSNLQILKELALSTSV
DVDPQEGEITIPMPLGKTEKELDMATPKELISEHPGEDSTFETIEQLDLRQ
ATSPERDAGLDLAAYSQLCFDLRFSPSRFGLLFHMYDSLTVPAQORALAF
EKGSVLFNIGALHTQIGARQDCSTEGTNHAAEFORAAAGAFLLRENTFSAFSDMS
AASLMLEOLMTAQOECIFKGLLPASATPDICPDQLOLAQAAVATRYGVHVRAM
AOPVVDYLPASWTNLAHVKAHEFCALAHVHAAMALCESHPAKGELAROBHVQSTP
HEPLGDTLPQHPEDRRKKAHLKAILQBEALRLHLTLRVLKVDLLQVVTQALR
RSATKLSQLEREDDFEATEADIPQKTHQTPGEPISVFNKRWHLVGVHVRMBG
GFGLTLRGDSFVLIAAVPGQASAGLKEGDIYVSNVQPCQKMWKHLVYVTLRSG
EKGSLQVVSLLPSPEPRGTGPRRAALLWNQRCGFPETPMTPTRTPTWILGWSRKNQ
GKTGSHDPCTNENCVCP"

ORIGIN

Query Match 44.6%; Score 1102.2; DB 10; Length 1932;
Best Local Similarity 75.9%; Pred. No. 1e-171;
Matches 1450; Conservative 0; Mismatches 383; Indels 78; Gaps 4;

QY 321 GGGCTGTGACTCCCTGACGCGAGATCCAGTCCGCGCCAGCTGCAGAGCGCGAGGCCACAGAT 380
DB 81 GGGCTGTGCTTCTGTGTGCAAAACAGCTCGCGCAGCTGCAGAGTCAAGGGCTCGGCT 140
QY 381 TCACACGAGCATGGAACAAGGAGCTGCAGATGCGGACCGGCGCTGAGAACCTCTACAGAGC 440
DB 141 CCACGAGCAGTCAACGAGGAGCTCCGATGCGGACCGGCTGCTGAAACCTCTACAGAGC 200
QY 441 CACGAGCAACACCGGTGAGAGAGCGTGGCGCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 500
DB 201 CACGAGCAACACCTGGGTCCGAGAAAACAGTCCGACTGAGAGTGAAGTGAAGTGAAGTGAAG 260
QY 501 CTTGCGAGCTCTGAGAGGAGCTGAGAGGAGCTCAGCGGTGGCGTGGCGTGGCGTGGCGTGGCG 560
DB 261 CTTGCGAGCTCTGAGAGGAGCTGAGAGGAGCTCAGCGGTGGCGTGGCGTGGCGTGGCGTGGCG 320
QY 561 TGGGAGCGAAGCTGTCTACTGTCCCCATGATCCCCCTGGGCTGAAGGAGGAGGAGGAGGAGGAG 620
DB 321 AGAGGGGAGGAGGAGTACTATTCCTATGATCCCCCTGGGCTGAAGGAGGAGGAGGAGGAGGAG 380
QY 621 GGAAGTGTCTACCGCTGAGAGGAGCTGATCTAGTGCACATTTGGAGAGGAGGAGGAGGAGGAG 680
DB 381 GGAAGTGTCTACCGCTGAGAGGAGCTGATCTAGTGCACATTTGGAGAGGAGGAGGAGGAGGAG 440
QY 681 CTAGGAG 740
DB 441 CTTTGAAGAGAGATCCAGGAACCTAGAAGACCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 500
QY 741 GAATGAGTGGGCGCTGAGAGTGTCTACAGGCTATTACACAGCTGTGCTTCTTCTTCTTCTTCTT 800
DB 501 GGAGGAGGCGGCGCTGATCTGCTCGCAGGCTACTATAGCAGCTCTGTTTCTTCTTCTTCTTCTT 560
QY 801 GCGCTTCT 860
DB 561 ACCTTTCT 620
QY 861 TGGGCTTCT 920
DB 621 GGGGCTTCT 680

QY 921 CGGTGCTCCACACGAGATTGGGGCGGCAGGACCGCTCTGACCGAGGGTGCCGG 980
Db 681 TGGGGCACTCCACACGAGATTGAGCAGCGGAGGATTGCTCTGTCACCGAGGGCACCAA 740
QY 981 CGCGCTATGAGAGGCTTCCAGAGGGCGCTGGGGCTTTAGCGCTCTCTGAGGGAGAACTT 1040
Db 741 CCACGCTGTGAGGCTTTTCAGAGGGCTGCTGGGGCTTCAGACTCTCTGAGGGAGAACTT 800
QY 1041 CTCCTATGCGCGAGCCCAAGACATGAGCGCTGCGTCCCTCTGCGCACTGAGCAGCTCAT 1100
Db 801 CTCCTATGCGCGAGCCCAAGACATGAGCTGCTGCTCTCTCTCATCTGAGCAGCTCAT 860
QY 1101 GATGCGCCAGGCGCCAGGAATGTTGTTTTCAGGGCTCTCACACCTGCTCATGCGCCCG 1160
Db 861 GATGCGCCAGGCGCCAGGAATGCTCTTCAAGGGCTCTTACTGCGCAGCTCTGCCACACC 920
QY 1161 CCAAGACTCTGCTGGCCCGCTGCGCTGCGCAGAGGGCGCGCCAGGTGGCAGCGAGTA 1220
Db 921 TGACATCTCTGACATCACTGACGTGGCCAGGAGCTGCGCAGGTGGCACTGAATA 980
QY 1221 CAGGCTAGTGCACCGGACCATGCGCCAGCACCCGCTCCAGCACTAGTGCCTGTCTCTG 1280
Db 981 TGGGCTTGTGCAACCGGCGCATGCGCCAGCACCTGTCCGAGACTACTCTGCTGCTCTG 1040
QY 1281 GACTGCGCTGTGTCATGTCAGGCGGAGTACTTCCGCTCCCTGCGCCCACTACAGTAGC 1340
Db 1041 GACTAACCTTGGCGCAGCTCAAGGCTGAGCATTCTGCGCGCTGCGCCCACTATCATGCA 1100
QY 1341 CATGCGCTCTGCGAGCGCTCCCGCAGCACCGAGGAGAGTCCCCACACGACGAGCT 1400
Db 1101 CATGCGCTCTGCGAAGCGCACCCAG---CAGAGGTGAACTGCGACGAGGACGCT 1157
QY 1401 CTCTCTGAGCGCCCGCACTCTCTTAAGCCCGAGGCGCTCTGCTGCGCAGGAGCTGGA 1460
Db 1158 CTTC---CAGCGCTCTCACTCCCGCAGCGCTGCGCGCCCGCACACTGCGCAGCACCC 1214
QY 1461 GGAGCGAGGAGCTTGGCAGGCACTGAGGAGGTCACCTGAGGAGGTCACCTGGGCGAGG 1520
Db 1215 AGCCGCGAGGAGCTTGGCAGGCGCACCTGAAGCGTGCATCTCTGCGCGCAGGAGGAG 1274
QY 1521 GCTGCGGCTGCAACGCGCTCTGCGCGCTGCTGCGCGAGGAGGAGCTCTGCGGCTGTGAT 1580
Db 1275 GCTGCGGCTTACACTTGTGAGAGTCTGCGCAGGAGTGGACCTGCTGCGAGTGTGCT 1334
QY 1581 CTCGAGAGCTGAGCGCTCACTGCGCAAGTATGCGGAGCTGACCGTGGAGTGAATGACT 1640
Db 1335 GACTCAGGCACTGCGCGCGCTCTCTGCGCAAGTACTCAGAGCTCGAGCGCGAGGATGACT 1394
QY 1641 CTGTGAGGCTGCGGAGGCGCGGACATCCAGCTTAAGCCACACGAGAGCCAGAGGCGCAG 1700
Db 1395 CTTGAGGCTACGAGGCGCGCGAGCATCCAGCTTAAGCTTCCAGAGCTTCCAGAGCCCGA 1448
QY 1701 GATGCCACGCTGTCCAGGGAAGGGGCTGACATCTTCATCGCTGGGCGCCCTGTCTC 1760
Db 1449 -----GGGACCTTGTCT 1460
QY 1761 TGTGTTCTCAGCAAGAACCGGTGCGGCTGTTGGGCGCGTCCACTGACCCGAGGAGA 1820
Db 1461 TGTGTTTTCACCAAGAACCGATGTCAGTGGTGGTGGCGGCTGACATGACTCGAGGAGA 1520
QY 1821 GGGCGCTTGGCTTGGCTTGGGAGACTGCGCTCTCTCATCGCTGCGCTCATTTCC 1880
Db 1521 GGGAGCTTGGGTTTTCAGCTGCGGAGAGACTCTCCAGCTCTCATCGCTGCTGCTGCC 1580
QY 1881 AGGAGCCAGGCGCGGCGCTTGAAGAGGCGGAGTACATTTGTGTCAGTGAATGG 1940
Db 1581 TGGGGTTCAGGCTGAGTTCGCTGGCTTGAAGAGGCGGAGTACATCTGCTGCTGCTGCT 1640
QY 1941 CGAGCCTGAGGCTGAGGAGACACGCGGAGGTGTGACGAGGCTGAAGGCTGCGGAGAGA 2000
Db 1641 GCAGCGGTGCAAGTGTGGAAGCACCTGAGGCTGTGTGACAGCTGAGGAGCATGGCGGA 1700
QY 2001 GCGCGGCGCCAGCTGCAAGTGTGTGCTGCTGCTGCCAGCTCTAGACTGCCAGCTGGG 2060

Db 1701 GGAAGCGGTGAGCTGCGAGTGTGTCTATTGTCGCCAGCCCTGAGCCAGTGCACAG 1760
QY 2061 GGACCGCGCGCGCTGCTGCTGCGGCGCCAGGCGGCTTCTAAGAGGAGCCAGGAGCATGG 2120
Db 1761 ACCCGAGCGGCGAGCACTGCTG-----TGGAACGAGAGAGTGTGG 1802
QY 2121 TTGCAAGACCCCGCATCCACGTGGCGCCAGTCCCGGCGCCCTCTCAACTGGAGCCGAAA 2180
Db 1803 TTTTGAGACACGATGTCGCCACAGTACCGCGCTGCGCCCTCTGCTGCTGCGGCGCGAA 1862
QY 2181 GSCCCAGAGGCGAAGACTGGAGGCTGCCCCAGAGCCCTGTGCCCCAGTGA 2231
Db 1863 GAATAACAAGGCAAGACTTGAAGCCACCTGACCCCTGCACCAACAGGAA 1913

BC052010 2301 bp mRNA linear ROD 08-OCT-2003
Mus musculus rhophilin, Rho GTPase binding protein 1, mRNA (cDNA
clone MGC:62337 IMAGE:5717786), complete cds.

BC052010
ACCESSION BC052010.1 GI:30354356
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

RESULT 8
BC052010
LOCUS
DEFINITION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

1 (bases 1 to 2301)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Uscin,I.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettner,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalish,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932

2 (bases 1 to 2301)
Strausberg,R.
Direct Submission
Submitted (01-MAY-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
cDNA Library Preparation: M. Bento Soares, University of Iowa
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILINL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 114 Row: k Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6680084.

FEATURES

source

1. .2301
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:62337 IMAGE:5717786"
/tissue_type="Brain, mouse 12.5 dpc"
/clone_lib="NIH BMAP_FDO"
/lab_host="DH10B"
/notes="Vector: pYX-ASC"

gene

1. .2301
/gene="Rhpn1"
/note="synonym: Rhophilin"
/db_xref="LocusID:14787"
/db_xref="MGI:1098783"

CDS

66. .1421
/codon_start=1
/product="Rhpn1 protein"
/protein_id="AAH52010.1"
/db_xref="GI:30354357"
/db_xref="LocusID:14787"
/translation="MIIEERPDGGTGEBSRRPDDGSIIRKYGVSFVNQPGQLQSHR
ARLHQISRLRMTVRENTWRETSVNTWRETSVNTWRETSVNTWRETSVNTWRETSV
DVDPGEGETTIIMPLGKETKELDWATPELKIISHFEDGTSFETFEQLELDLRQ
ATTPRDEAGLDMIAAYSQLCFDARFSPRSPLGLFHWYDSLTGVPAQORALAF
EKSGLFNIGALHTOIGARODCSCTECTNHAARAFORAAAGAFRLIRNFHAPSDDMS
AASLMLEQIMIAQARCEIFKGLLLPASATPDICPDQLQAQEAQVATYEYGLVHRAM
AQPPRDYLPASNTLAAHVAEHFALAHYAHAALCEISHPAKGELGARQEHVFPSTP
HEPLGPTLPQHPEDRRKLGGGAASHLVQSPAQGGPAAGCGDSGTALPQGVILTARAR
G"

misc_feature

189. .404
/note="HR1; Region: Hrl repeat"
/db_xref="CDD:pfam02185"

misc_feature

408. .966
/note="BR01; Region: BR01-like domain. This functionally
uncharacterised domain is found in a number of signal
transduction proteins, including Rhophilin and BR01"
/db_xref="CDD:pfam03097"

ORIGIN

Query Match 42.3%; Score 1043.8; DB 10; Length 2301;
Best Local Similarity 70.0%; Pred. No. 4.1e-162;
Matches 1550; Conservative 0; Mismatches 522; Indels 143; Gaps 5;

Qy	321	GGGCTGTGACTCCCTGACGAGATCCAGTGGCGCAGCTGCAGAGCGCGAGGCCAGAT	380
Db	146	GGGCTATGGCTCCCTTTGTGAAACCAGCTGGCCAGCTGCAGAGTCAAGGGCTCGCT	205
Qy	381	TCACGACGAGATTGACAAAGAGCTGCAGATGCGGACGGGCGCTGAGAACCTTACAGAGC	440
Db	206	CCACCAGAGATCAGCAAGAGGCTCCGAGTGGGACCGGTGCTGAAACCTTACAGAGC	265
Qy	441	CACGAGCAACACCGGTGAGAGAGACGGTGGCCCTGGAGCTGAGCTACGTCACTCAA	500
Db	266	CACCAGCAACACCTGGGTCGCGAAGAACAGTGCACCTGGAGCTACGTCACTCAA	325
Qy	501	CCTGCAGCTGCTGAAGGAGGCTGGAGGAGCTCAGCGGTGGCGTGGACCTGGCCGCA	560
Db	326	CCTGCAGCTGCTGAAGGAGGCTGGAGGAGCTCAGCAGCTGAGCACCGCTAGTGTGGACAGCC	385
Qy	561	TGGAGGCAAGCTGTCTACTGTCCCATGATCCCTGGGCTGAAGAGACCAAGGAGCT	620
Db	386	AGAGGGCAAGGATTACTATTCCATGATCCCCCTGGGCTGAAGGAAACCAAGGAGTT	445
Qy	621	GGACTGGTCTACACCGCTGAAGGAGCTGATCTCAGTGACATTTTGGAGGAGCGGCTC	680
Db	446	GGACTGGGCCACACCGCTGAAGGAGCTGATCTCTGAGCACTTTTGGAGGAGTGTACTCTC	505

Qy	681	CTACGAGGAGAGAAATCAGGGAGCTGGAGGCCCTGGCGAGGCCATCGGACCCCGACCGC	740
Db	506	CTTTGAGACAGAGATTCAGGAATCAGGAACCTACGCGAGGCCACGAGGACTCCAGCAG	565
Qy	741	GAATGAGTGGGGCTGGAGCTGCTCAGCCTATTACAAACAGCAGTGTCTCTTGATGC	800
Db	566	GGACGAGGGGGCTGGATCTGCTCGCAGCTACTATAGCAGCTCTGTCTTCTGGACGC	625
Qy	801	GGCTTCTCTACACCTTCGCCAGAGGCTCGGGCTCTTCTTCATCTGTTACGACTCGCTAC	860
Db	626	ACGTTTCTTCAGTCCCTCCAGGAGCCCGGGCTCTTCCATCTGGTACGACTCGCTAC	685
Qy	861	TGGGGTCCCGGCCACGAGCGTCCCTGGCTTCGAGAAGGCGAGCGCTTCTCTTCAACAT	920
Db	686	GGGGTCCCGGCCACGAGCGGCTCTGGCTTCGAGAAGGCGAGTACTCTGTTTACAT	745
Qy	921	CGGTGCTTCCACACGAGATTGGGGCGGCCAGGACCGCTCTCTGCAACGAGGGTGC	980
Db	746	TGGGGCACTCCACACGAGATTGGAGCACGCGAGATTTGCTCTCTGCAACGAGGGCAA	805
Qy	981	CCGGCTATGGAGGCTTCCAGAGGCGCTGGGGCTTTCAGCTCTCTGAGGAGAACTT	1040
Db	806	CCACGCTGTGAGGCTTTCAGAGGGCTGCTGGGGCTTTCAGACTCTCTGAGGAGAACTT	865
Qy	1041	CTCCCTGCGCCGAGCCAGACATGAGCGCTGCTCTGCGCACTGCGAGCAGCTCAT	1100
Db	866	CTCCACGCGCCAGGCCAGACATGAGTGTGCTCTCTCTCTCCATGCTGAGCAGCTCAT	925
Qy	1101	GATGGCCAGGCCAGGAATGTGTGTTGAGGCGCTCTCAACCTGCTCTCATGTCCTG	1160
Db	926	GATTGCCAGGCAAGGAATGCACTTCTCAAGGGCTTTTACTGCGACCTCTGCGACACC	985
Qy	1161	CCAAGACTGCTGGCCAGCTGCGCTGGCGGAGGCGCCCGAGTGGCGAGCGAGTGA	1220
Db	986	TGACATCTGTCCAGATCACTGAGCTGGCCCGAGGAGCTGCCAGGTGGCACTGAATA	1045
Qy	1221	CAGCTAGTGCACCGGACCATGCGCCAGCCACCGCTCCAGACTACGCTGCTCTCTG	1280
Db	1046	TGGGCTGTGCAACGGGCGCATGGCCAGCCACCTGTCCGAGACTACCTGCTGCTCTG	1105
Qy	1281	GACTGCGCTGGTGTGATGTAAGGCGGAGTACTTCCGCTCCCTGCGCCACTACGAGTAC	1340
Db	1106	GACTTAACCTGGCGCAGCTCAAGGCTGAGCACTTCTGCGCCCTGCGCCACTATCATG	1165
Qy	1341	CATGGCCCTCTGCGAGCGCTCCCGAGCGACCGGAGGAGAGCTCCCGACCGACGAGG	1400
Db	1166	CATGGCCCTCTGCGAAGACCAACCCAG---CAAGGGTGAAGTGCAGAGGAGGAGCT	1222
Qy	1401	CTTCTGTCAGCCCCCACTCTCTTAAGCCCCGAGGCGCTGTGCTGCGCGAGGAGTGA	1460
Db	1223	CTTC---CAGCCCTCCACTCCCATGAGCCCTGGGCCCCACACTGCGCGAGCACCAGA	1279
Qy	1461	GGAGCGGAGGAGCTTGGCAAGGCAACCTGAAGCGTGCCATCTCTGGGGCAGGAGGAGC	1520
Db	1280	AGACCGCAGGAAGCTT-----GGAGGAGGC 1304	
Qy	1521	GCTGCGCTGACGCGCTGTGCGCGCTCTGCGGAGGTGACCTGCTCGGGCTGTGAT	1580
Db	1305	GCTGCGCTTCAACACTGTGCGAGTCTGCGCAAGGTGGACCTGCTCAGGTTGTGTT	1364
Qy	1581	CTCCAGACGCTGCGAGCTCACTGGCCAAAGTATGCGAGCTGCACCGTGAAGATGACTT	1640
Db	1365	GACTCAGGCACTGCGGCGCTCCCTGGCCAAAGTACTCAAGCTTCAGCGCGGAGTGA	1424
Qy	1641	CTGTGAGGCTGCGGAGGCGCGGACATCCAGCTTAAGACCCACAGAGCCAGAGGCCAG	1700
Db	1425	CTTCAGAGGCTACGAGGCGCGGACATCCAAACCCAGACTCACAGACCCAGAGGTGAG	1484
Qy	1701	GATGCCAGGCTGTCCAGGGGAAGGGGCTTGAATCTTTCATCGGCTGGGGCCCTGTC	1760
Db	1485	GATGCCCTAGCTGTGCGAGGTGAAGTGAAGTCTTCTTCACTCGGCTGGGACCTTGT	1544
Qy	1761	TGTGTTCTCAGCCAAAGAACCGGCTGGCGGCTGTTGGGGCCCGTCCACTGACCCGAGGAGA	1820


```
Db 1545 TGTGTTTTCACCAAGAACCGATGGCAGTTGGTAGGCGCGGTGCACATGACTCGAGGAGA 1604
QY 1821 GGGCGGCTTTGGCTCAGCTTCGCGGAGACTCGCTGTCCTCATGCTGCGCTCATTTCC 1880
Db 1605 GGGAAAGCTTCGGGTTTCAAGCTGCGGAGACTCTCCAGTCTCTCATGCTGCTGCTGCC 1664
QY 1881 AGGGAGCCAGGCGCGG----- 1897
Db 1665 TGGGGGTGAGGCTGAGGTAAAGCCCTCTAGCCCGGGTGTGGAATCCTAGAGGACCGC 1724
QY 1898 -----CGGCTGGCTGAAGGAGG 1916
Db 1725 TAGATACTGCCACACCTCACCGATCCACATCTCCGCGATCCGCTGGCTGAAGAGGG 1784
QY 1917 CGACTACATTTGTCAGTGAATGGGAGCCTATGCGAGTGTGGTGGAGACACGCGAGGTGTG 1976
Db 1785 CGACTACATCTGTCGCTGCAACCGGAGCGCTGCAAGTGTGGTGAAGCACCTGGAGTGTG 1844
QY 1977 GACGAGCTGAGGCTGCGGAGAGCGGGGCGCAGCTTGCAGTGTGGTGTGCTGCTGCC 2036
Db 1845 GACACAGCTGAGAGCATGCGGAGAGCGGTGAGCCTGCGAGTGTGCTGCTGCTGCC 1904
QY 2037 CAGCTCTAGACTGCCAGCTTGGGAGACCGCGCGCGCTGCTGCTGGGCGCCAGGGGCT 2096
Db 1905 CAGCCTGAGCCACGTGGSCACAGGACCGCGAGCGGAGCAGCTGCTG----- 1950
QY 2097 TCTAAGAGCCAGAGGAGCATGTTGCAAGACCCCGCATCCAGCTGGGCGAGTCCCG 2156
Db 1951 ----TGGAAACAGAGAGTGTGGTTTGGAGACCGATGCCACACGTACCGCGCCCTG 2006
QY 2157 GCCCTCTCACTGGAGCCGAAAGCCGAGGCGCAAGACTGAGGCTGCGCCCGACCC 2216
Db 2007 GCCCATCTTGGCTGGAGCCGCAAGATAAACAAGGCAAGACTGGAAGCCACCTGACCC 2066
QY 2217 CTGTGCCCCAGTGAAGCAGCTCCCGCTTCATCTTGAAGACCCAGGCTGCGCTGAGG 2276
Db 2067 CTGCACCAACAGGACTGTGTACCTGCCCCCTGAGCTGGACACCCCGTATCCCTGTGTG 2126
QY 2277 GCGAGGATCCTGACGCGCTCAGCCCTGAGCTTCCAGCTGGCAGCAGCAGCAGCATGCC 2336
Db 2127 TCAGGATGGCTGTGGGGGCCATCTCTATGTATCATCACCTTGAATCAGCTCCAACTAG 2186
QY 2337 CTCCCCACCCAGAGACCTCCGGGAATGCTGCTCCGCTCATGCTGAGGCTGCTGCTG 2396
Db 2187 CGAGATGCTCTAGCGGGTCTGTCTTACTCTTGGGGGACTATCAGGCTGCGCCACCT 2246
QY 2397 GGCACCTGCCCTGATTAAGACTGGTTCAGACCTGCTGTAAGAAAAA 2451
Db 2247 TGCTTAGAGATGTCTATTAAAAATGCTGGTCAGATGAAAAA 2301
```

```
RESULT 9
BD183400
LOCUS BD183400 4729 bp DNA linear PAT 17-JUN-2003
DEFINITION Novel genes and proteins encoded by the genes.
ACCESSION BD183400
VERSION BD183400.1 GI:31875600
KEYWORDS JP 2002345492-A/113.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 4729)
Ohara,O., Nagase,T. and Nakajima,D.
Novel genes and proteins encoded by the genes
Patent: JP 2002345492-A 113 03-DEC-2002;
KAZUSA DNA RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002345492-A/113
PD 03-DEC-2002
PF 26-FEB-2002 JP 2002049009
PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
```

```
PC C12N15/09, C07K14/47//A61K31/711, A61K38/00, A61K48/00, A61E25/00,
PC A61P25/14,
PC A61P25/18, A61P35/00, C12N15/00, A61K37/02
CC Novel genes and proteins encoded by the genes FH Key
Location/Qualifiers
FT CDS (518)..(1540).
FEATURES
source Location/Qualifiers
1..4729
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
```

ORIGIN

```
Query Match 38.6%; Score 952.8; DB 6; Length 4729;
Best Local Similarity 93.2%; Pred. No. 3.5e-147;
Matches 1018; Conservative 0; Mismatches 72; Indels 2; Gaps 2;

QY 1346 CCCTCTGCGAGCGCTCCCGAGGACCGGAGGAGAGCTCCCGACGACGAGCGAGTCTTCC 1405
Db 2235 CCGCGTGGCGGCACCCCGCAAGAAAGTGGCTGTGATGAGCCCCACAGCCCTGGCGTTGC 2294
QY 1406 TGCAGCCCCCACCCTCTTAAGCCCCGAGGCCCTGTGCTGCCGAGGAGCTGAGGAGC 1465
Db 2295 CCATCTCTTCTGCCAGCTCCCGAGGCCCAATGTTGTGACATCCCAAGTGC 2354
QY 1466 GAGCAGCTTTGGCAAGGACACCTGAAGCGTGCATCTTGGGGCAGGAGGCGCTGC 1525
Db 2355 CCGCGCTGC-AGGCAAGGACACACCTGAAGCGTGCATCTTGGGGCAGGAGGCGCTGC 2413
QY 1526 GCTGTCAGCGCTTGGCGCGCTCTGCGCGAGGTGGACCTGCTTGGGGTGTGATCTCC 1585
Db 2414 GCTGTCAGCGCTTGGCGCGCTCTGCGCGAGGTGGACCTGCTTGGGGTGTGATCTCC 2473
QY 1586 AGCGCTCAGCGCTCACTGCGCAAGTATGCGGAGCTCGACCGTGAGGATGACTTCTGTG 1645
Db 2474 AGCGCTCAGCGCTCACTGCGCAAGTATGCGGAGCTCGACCGTGAGGATGACTTCTGTG 2533
QY 1646 AGGTGCGCGAGGCCCGGACATCCAGCTTAAGACCCAGAGCCAGAGGCGCAGGATGC 1705
Db 2534 AGGTGCGCGAGGCCCGGACATCCAGCTTAAGACCCAGAGCCAGAGGCGCAGGATGC 2593
QY 1706 CAGCGCTTCCCGAGGAGAGGGCTGACATCTTCATCGGCTGGGGGCCCTGTCTGTGT 1765
Db 2594 CAGCGCTTCCCGAGGAGAGGGCTGACATCTTCATCGGCTGGGGGCCCTGTCTGTGT 2653
QY 1766 TCTCAGCAAGAAACCGGTGGCGCTGGTGGGCGCTGACATCTTCATCGGCTGGGGGCC 1825
Db 2654 TCTCAGCAAGAAACCGGTGGCGCTGGTGGGCGCTGACATCTTCATCGGAGAGGGCG 2713
QY 1826 GCTTTGGCTCACGCTTGGGGAGACTCGCTGTCTCATCGCTGCGCTCATTCAGGGA 1885
Db 2714 GCTTTGGCTCACGCTTGGGGAGACTCGCTGTCTCATCGCTGCGCTCATTCAGGGA 2773
QY 1886 GCGAGGCGCGCGCTGGCTTGAAGAGGCGGACTACATTTGTGTGATGATGGGAGC 1945
Db 2774 GCGAGGCGCGCGCTGGCTTGAAGAGGCGGACTACATTTGTGTGATGATGGGAGC 2833
QY 1946 CATGCAAGTGTGGAGACACCGGAGGTGGTGGAGCTGAAGCTGCGGAGAGGGCGG 2005
Db 2834 CATGCAAGTGTGGAGACACCGGAGGTGGTGGAGCTGAAGCTGCGGAGAGGGCGG 2893
QY 2006 GCGCCAGCTGCAAGTGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACC 2065
Db 2894 GCGCCAGCTGCAAGTGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACC 2953
QY 2066 GCGCCAGCTGCTGCTGCGGCGCCAGGGGCTTTAAGAGCGCAGAGGAGCATGGTTGCA 2125
Db 2954 GCGCCAGCTGCTGCTGCGGCGCCAGGGGCTTTAAGAGCGCAGAGGAGCATGGTTGCA 3013
QY 2126 AGACCCCGGATCCACGCTGGGCCAGTCCCGGGCCCTCTCACTGAGCGCGGAGGCGCC 2185
Db 3014 AGACCCCGGATCCACGCTGGGCCAGTCCCGGGCCCTCTCACTGAGCGCGGAGGCGCC 3073
```



```

RESULT 12
BC036447
LOCUS
DEFINITION
Homo sapiens rhophilin, Rho GTPase binding protein 2, mRNA (cDNA
clone MGC:26769 IMAGE:4830913), complete cds.
ACCESSION
BC036447
VERSION
BC036447.2
KEYWORDS
GI:34784999
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1. (bases 1 to 3516)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS
Srausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Vallalun, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2238257
PUBMED
12477932
REFERENCE
2 (bases 1 to 3516)
Srausberg, R.
Direct Submission
Submitted (09-AUG-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
On Sep 16, 2003 this sequence version replaced gi:22209018.
Contact: MGC help desk
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbiology.org
contact: amadansystemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketterman, Anuradha
Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAP plate: 34 Row: j Column: 22
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 23510298.
Location/Qualifiers
1. .3516
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:26769 IMAGE:4830913"
/cisue_type="Testis"
/clone_lib="NIH_MGC_"
/lab_host="DH10B"
FEATURES
source
1. .3516
/note="Vector: pBluescript"
/gene="RHPN2"
/note="synonym: p76RBE"
/db_xref="LocusID:85415"
37. .2097
/codon_start=1
/product="rhophilin-like protein"
/protein_id="AAH36447.1"
/db_xref="GI:22209019"
/translation="MTDALLPAAPQPLEKNDGYFRKGNPLAQTRGSKLQNRQAALN
QQILKAVRMTGAENLLKVATNSKVRQVLELSFVNSDIQMLEEGLNLSVGVYQ
NTEAEFTPLIPLGKETQDVFAVVKDFILEHYSDEGGLYVEDEIADMLDLQACRT
PSDEAGVELLMTYFIQLGVRSEFPPTROMGLLFTWYDSLTVGPVSQONLLLEKAS
VLNTEALYIQICRCDRQTOAGLESNAIDAFORAAVNLKDTFTHTPSYDMSAPML
SVLVKMLAQAESVFEKISLPGRNEFPMVKVQAEAAKGVYQQLHAAMSQAPVK
ENTPYWSALACVKAHYAALAHYFTAILLDHVKPGLDHOEKCLSLQYDHMPG
LTPLATKNDQRRQLGKHLRMAHHEESVREASLCKKLSIEVLQVLCQAQERS
RTTYAQHQBEDLLNLDAPSVAKTEQEVDIILPQFSKLTVDTPFKLGLPVSFSAN
KWTTPRSIRFTAEGDGLGFTLRGNAPVQVHFLDPYCSASVAGAREGYIVSLVDC
KWLTLSEVMKLLKSFGEDETEMKVSLDSTSMHNKSATYSVGMOKTYGMICLAIDD
DKTKDKIKSKLSFLSGTNRKOKSASTLCLPSVGAARPQVKKLPSFPFSLNSD
SSWY"
163. .354
/note="Hr1; Region: Hr1 repeat"
/db_xref="CDD:pfam02185"
367. 825
/note="BR01; Region: BR01-like domain. This functionally
uncharacterised domain is found in a number of signal
transduction proteins, including Rhoophilin and BR01"
/db_xref="CDD:pfam03097"
1576. 1809
/note="PD2; Region: Domain present in PSD-95, Dlg, and
ZO-1/2. Also called DHR (Dlg homologous region) or GLGF
(relatively well conserved tetrapeptide in these domains).
Some PDZs have been shown to bind C-terminal polypeptides"
/db_xref="CDD:smart00228"
ORIGIN
Query Match 20.1%; Score 496.4; DB 9; Length 3516;
Best Local Similarity 56.8%; Pred. No. 5.2e-72;
Matches 977; Conservative 0; Mismatches 731; Indels 12; Gaps 3;
Qy 321 GGGCTGTGACTCCCTGACCGAGATCCAGTGGCGGCGAGTCGACGCGCGAGGCCAGAT 380
Db 105 GGGCTGTGTAATCCCTTGCACAAACCGCGCGAGTAATTCAGAAATCAAGAGCTGCTT 164
Qy 381 TCACCGACAGATTGACAAGGAGCTGCGAGATCGGACCGGCGCTGAGAACCTCTACAGAGC 440
Db 165 GAATCAGCAGATCTCGAAAGCCGTGCGGATGAGGACCGGAGCGGAAACCTTCTGAAAGT 224
Qy 441 CACCAGCAACACCGGCTGAGAGACGCTGCCCTGAGCTGAGCTACGCTCAACTCAA 500
Db 225 GGCCACAAACTCAAAGGTGCGGAGCAAGTGGCGCTGAGCTGAGCTTCGCTCAACTCAGA 284
Qy 501 CCTGCAGCTGCTGAAGCAGGAGCTGGAGGAGCTCAGCGGTGGCGTGGACCCCTGGCCGCA 560
Db 285 CCTGCAGATGCTCAAGGAGAGCTGGAGGGGCTGAACATCTCGGTGGCGCTTATCAGAA 344
Qy 561 TGGGAGGAGAGCTGCTGCTGCTCCATGATCCCTGGGCTTGAAGAGACAGGAGCT 620
Db 345 CACAGAGGAGGAGCATTTACGATTCCTCCCTGATTCCTTGGGCTGAGGAAACGAAAGAGT 404
Qy 621 GGACTGCTGTACACCGCTGAAGGAGCTGATCTCAGTGCACCTTTCGAGGAGGACGCGCTC 680
Db 405 CGACTTTGAGTCTCTCAAGGATTTTATCTGGAACATTTACAGTGAAGATGCTATTT 464
Qy 681 CTACGAGGAGGAGAAATCAGGAGGAGCTGGAGGCCCTCGCGCAGGCCATGCGGACCCGCG 740
Db 465 ATATGAAGATGAAATTCAGATCTTATGATCTGAGACAAAGCTTGTTCGAGCGCTAGCG 524
Qy 741 GAATGAGTGGGCGCTGGAGCTGCTCAGACGCTATTACACAGCTGTGCTTCTCTGATGC 800

```

Db 525 GATGAGGCGGGGTGGAACTCTGATGACATCTTCATCCAGCTGGGCTTTGTGAGAG 584
QY 801 GCGCTTCTCCATCCCTGCGAGGAGCTCGGGCTCTTCTTCCACTGTGACAGCTCGCTTAC 860
Db 585 TCGATTCTTCCCGCCACACGCGAGATGGACTCTCTGTTTCACTGATGACTCTCTCAC 644
QY 861 TGGGGTCCCGCCACGAGCGTGGCTTCCGCTTCCAGAAAGGAGAGGTTCTCTTCAACAT 920
Db 645 CGGGGTTCGGGTFCAGCCAGCAACCTGCTGTGGAGAAAGGCGAGTGTCTCTGTTCACAC 704
QY 921 CGGTGCGCTTCCACACGACATTTGGGCGCGCGCAGAGACCGCTCTCCACGAGGGTGGCCG 980
Db 705 TGGGGCGCTTCTACACCGAGTTGGGACCGGTGTGATCGGAGAGCGAGGTGGGTGGA 764
QY 981 CGCGCTATGGAGGCTTCCAGAGGCGCGCTGGGGCTTTCAGCGCTCTCGAGGAGAACTT 1040
Db 765 GAGTGCCATAGATGCTTTCAGAGAGCGCGAGGGTTTTAAATTACCTGAAGACACATT 824
QY 1041 CTCCTATGGCGGAGGCGAGATGAGCGCTGGCTCGCTTCTGGCACTGGAGAGCTCAT 1100
Db 825 TACCCATATCTCAAGTTTACGACATGAGCGCTGCCATGCTCAGCGTGTCTGTCAAAATGAT 884
QY 1101 GATGCGCCAGGCGCCAGGAATGTGTTTGGGCGCTCTCACCACTGCTCCATGCGCC 1160
Db 885 GCTTGCAACAGCCCAAGAAAGCGTGTGGAGAAATCAGCCTTCTCTG-----GATCG 938
QY 1161 CAAAGACTGCTTGGCGCCAGCTGCGCTGCGCAGAGGCGCGCCAGGTGGCAGCCAGTA 1220
Db 939 GAATGAATTTCTATGCTGTGTAAGGTGCTCAGAGAGCTGTAAAGTGGGAGAGTCTA 998
QY 1221 CAGGTAGTGACCGGACCATGCGCCAGCCACCGTTCAGACTGAGTGGCTGTCTCTG 1280
Db 999 CAAACAGTACACGAGCGCATGAGCGCGCGGTGAAAGAGAAATCCCTACTCTCTG 1058
QY 1281 GACTGCGCTGTGTGATGTCAAGCGCGAGTACTTCCGCTCCCTGGCGCCACTACAGTAC 1340
Db 1059 GCGCAGCTTAGCTGCTGTAAGGCCCACTACAGCGGCGCTGGCCCACTACTTCACTGC 1118
QY 1341 CATGCGCTCTGCGAGCGCTCCCGACGACGAGGAGAGCTCCCGACGACGAGCAGT 1400
Db 1119 CATCTCTCTATCGACACACAGGTGAAGCGACCGATCTGGACACCGAGGAGAGTG 1178
QY 1401 CTCTCTGACG-----CCCCAGCTCTTAAGCCCGGAGGCGCTGTGC--TGCCGACGA 1454
Db 1179 CCTGTCCAGCTCTACGACCAATCCAGAGGGGCTGACACCCCTTGGCGCACACTGA 1238
QY 1455 GCTGAGGAGCGCAGCGAGCTTGGCAAGCGACACCTGAAGCGTGCATCTCGGGCAGGA 1514
Db 1239 TGATCAGCAGCGCCGACAGCTGGGAAGTCCCACTTGGCAGAGCCATGGCTCATCAGA 1298
QY 1515 GAGGCGCTGCGGCTGACGCGCTGTGCGGCTCTGCGGAGGTGAGCTGCTTCCGGCC 1574
Db 1299 GGAGTGGTGGGGAGGCGAGCTCTGCAAGAGCTGCGGAGCATTTAGGTGCTACAGA 1358
QY 1575 TGTGATCTCCAGAGCGTGCAGCGTCTACTGCGCAAGTATGCGGAGCTGCACCGTGAGA 1634
Db 1359 GGTGCTGTGTGCGGCACAGGAACCGTCCCGGCTACGTCACGCCACGACCGAGGAGGA 1418
QY 1635 TGACTTCTGTGAGGCTGCGGAGGCGCCGACATCCAGCTTAAGCCACCGAGAGCCAGA 1694
Db 1419 TGACTGCTGAACCTGATGAGCGCCGCCAGTGTGTTGCTGCTAACTGAGCAAGAGTTGA 1478
QY 1695 GGCCAGGATGCCAGCTGTCCAGGGGAAGGGGCTTGACATCTTCCATCGGCTGGGGCC 1754
Db 1479 CATTATATTGCCAGTTCTCAAGCTGACAGTACAGGACTTCTTCCAGAGAGCTGGGCC 1538
QY 1755 CTGTCTGTGTTCTCAGCCAGAAACCGGTGGCGGTGTGGGCGCCGCTCCACCTGACCCG 1814
Db 1539 CTTATCTGTGTTTTCGGGTAAACAGCGGTGAGCGCTCTCTCGAAGCATCCGCTTCACTGC 1598
QY 1815 AGGAGAGGCGGCTTTGGCTCAGCTTCGGGGAGACTGCGCTGTCTCTCATGCTGCGCT 1874

Db 1599 AGAAGAAGGGGACTTGGGGTTCCACTTGAGAGGGAACGCCCGCTTCAAGTTCACTTCT 1658
QY 1875 CATTCAGGAGGCCAGCGCGCTGGCTGCTGAAGGAGGCGGACTCATTTGTGTCACT 1934
Db 1659 GGATCTTACTCTGCTGCTGCTGCGGAGGAGCCCGGAGAGAGATTATTTGTCTCCAT 1718
QY 1935 GAATGGGAGCAGCATGCGAGGTGGTGAGACACCGGAGGTGGTGACGAGCTGAAGGCTGC 1994
Db 1719 TCAGCTTGTGATTGTAAGTGGCTGACGCTGAGTGAGGTATGAAGCTCTGAAGAGCTT 1778
QY 1995 GGGAGAGGCGGCGCGCAGCCTGACGCTGAGGTGGTGTGCTGCTG 2034
Db 1779 TGGCGAGGAGCAGATCGAGATGAAAGTCGTGAGCCTCTG 1818

RESULT 13

HSA347750

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

HSA347750 2061 bp mRNA linear PRI 06-JAN-2003
Homo sapiens mRNA for p76RBE protein.
AJ347750
AJ347750.1 GI:23304861
p76RBE protein.
Homo sapiens (human)
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Mitsescu, H., Steuve, S., Savonet, V., Degraef, C., Mellor, H.,
Dumont, J.E., Maenhaut, C. and Pirson, I.
Identification and characterization of a novel activated RhoB
binding protein containing a PDZ domain whose expression is
specifically modulated in thyroid cells by cAMP
Eur. J. Biochem. 269 (24), 6241-6249 (2002)

2
12473120
2 (bases 1 to 2061)
Direct Submission
Pirson, I.
Submitted (20-SEP-2001) Pirson I., Interdisciplinary Institute
IRIBHN, Université Libre de Bruxelles, 1070, Brussels, BELGIUM
Location/Qualifiers
1..2061
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..2061
/function="putative RhoB effector"
/codon_start=1
/product="p76RBE protein"
/protein_id="CAC87939.1"
/db_xref="GI:23304862"
/db_xref="GOA:Q8IUC4"
/translat="MTDALLPAAPLEKENDGYFRKGNPLAQTGRSKLQNRALN
QQILKAVRMRTGAENLLKVNATSKVQREVLSEFVNSDLQMLKEELNLSVGVYQ
NTEAFPTPLTGLKTDVDFVNLKDFLEHYSEDGLYDEIDADMLQACRT
PSRDEAGVELIMTYFIOLGFVSRFPPTROMGLFTWYDSLTPGVPSQOHLLEKAS
VLFTGALYTGITGRCDROTAQGLSAIDAFORAGVNLKDFTHPTSDVMSAPML
SVLVMKLAQESVFEKISLPGRNFFMLVKVAQEAKEVYQQLHAAMSQAPVK
ENIYFWSALCAKVAHYAALAHYFTAILLDHVKFGFTDLDHKECLSQLDYHMEG
LTPATLQNDQORQLGSHLRAMAHHEESVREASLCKLRSIEVLQVLCVAQERS
RLTYAQHEDDLNLIDAFSVVAKTEQVEDIILPOFSLKLTVDFFQKGLPSVFSAN
KRWTPPISIRFTAEDEGLFTLRGNAPQVHFLDPYCSASVAGAREGDIYVSIQLVDC
KWLFLSVMKLLKSGFDELEMKVVSILDSTSSMHNKSATVSGMOKTYSMICLAID
DDKTDKTKXISKSLSLFSLWGTNKNRQKRSASTLCLPSVGAARPVKKLPSFSLNSD
SSWT"

ORIGIN

Query Match

Best Local Similarity 20.0%; Score 494.8; DB 9; Length 2061;

Matches 976; Conservative 0; Mismatches 732; Indels 12; Gaps 3;

QY 321 GGGCTGTGACTCCCTGACGAGATCCAGTGGCGGACGCTGCAGAGCGCGAGGCCAGAT 380

```

1143 CCTGTCCAGCTCTACGACCACATGCGAGAGGGGCTGACACCCCTTGCGCACACTGAAGAA 1202
1455 GCTGGAGGAGCGCAGGAGCTTGGCAAGGCACACCTGAGCGTGCCATCTGCGGGCAGGA 1514
1203 TGATCAGCAGCGCGCAGAGCTGGGGAAGTCCCACTTGGCGAGAGCCATGCTCATCAGA 1262
1515 GGAGGGCGCTGCGGCTGCAAGCGCTGTGCGCGCTCTTGGCGAGGTGAGACCTGCTCGGGC 1574
1263 GGAATCGGTGGGAGGCGAGCTCTGCAAGAGCTGCGGAGCATTTAGGTGCTACAGAA 1322
1575 TGTGATCTCCAGAGCGCTGAGCGCTCACTGGCGCAAGTATGCGGAGCTGCACGTCAGGA 1634
1323 GGTGCTGTGTCGCGCAGAGAAAGCTCCCGCTCAGCTACGCGCAGCAGCAGGAGGA 1382
1635 TGAATTTCTGTGAGGCTGCGGAGCGCGGAGCATCCAGCTTAAGACCCAGAGAGCCAGA 1694
1383 TGACCTGTGTAACCTGATGACCGCCCCAGTGTGTTGCTTAAACTGAGCAAGAGTTGA 1442
1695 GGCGAGGATGCCAGCGCTGTCCAGGGGAAGGGCGCTGACATCTTCCATCGGCTGGGGCC 1754
1443 CATTATATTGCCCGCAGTTCTCCAAGCTGACAGTCACGGACTTCTTCCAGAAGCTGGGCC 1502
1755 CTGTCTGTGTTCTACGCGCAAGAACCGGTGGCGCTGTTGGGCGCGCTCCACCTGACCCG 1814
1503 CTTATCTGTGTTTTCGGCTAACCAAGCGGTGAGCGCTCTCTCGAAGCATCCGCTTCTG 1562
1815 AGGAGAGGGCGCTTTGGCTCTACGCTTGGGGGAGACTCGCTGTCTCTATCGTGCCTG 1874
1563 AGAAGAAGGGGACTTGGGGTTTCACTTGAAGGGAACGCCCGCTTTCAGGTTTCACTTCT 1622
1875 CATTCCAGGAGGAGCAGCGCGCGCTGCTGAAAGGAGGCGACTACATTTGTTCAGT 1934
1623 GGATCTCTACTGCTGCTCGCTCGGTGGCAGGAGCCGGAAGGAGATTATATTGCTCCAT 1682
1935 GAATGGCAGCCATGAGGTGGTGGAGACAGCGGAGGTGGTACGAGCTGAAGGCTGC 1994
1683 TCAGCTTGTGGATTGAAGTGGCTGACGCTGAGTGGATTATGAAGCTGCTGAAGAGCTT 1742
1995 GGAGAGGGCGCGCGCAGCGCTGCGAGTGTGTGCTGCTGCTG 2034
1743 TGGCGAGGACGAGATGAGATGAAGTCTGAGGCTTCTG 1782

RESULT 14
AX835273 3500 bp DNA linear PAT 15-DEC-2003
LOCUS
DEFINITION Sequence 2397 from Patent EPI347046.
ACCESSION AX835273
VERSION AX835273.1 GI:39921408
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and
Masuko,Y.
TITLE Full-length cDNA sequences
JOURNAL Patent: EP 1347046-A 2397 24-SEP-2003;
RESEARCH Association for Biotechnology (JP)
FEATURES
source 1..3500
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 20.0%; Score 494.8; DB 6; Length 3500;
Best Local Similarity 56.7%; Pred. No. 9.5e-72;
Matches 976; Conservative 0; Mismatches 732; Indels 12; Gaps 3;
```

```

69 GGGCTGTAATCCCTTGCACAAACCGCGCGAGTAATTCGAGATCAAGAGCTGCTTT 128
381 TCACCGAGAGATTGACAAGGAGTGTGAGATGCGGACGGGCGCTGAGAACTCTACAGAGC 440
129 GAATCAGCAGATCTGAAAGCCGTGGGATGAGGACCGGAGCGGAAACCTTCTGAAAGT 188
441 CACAGCAACAAACGGGTGAGAGAGCGTTCGCCCTGGAGCTGAGCTACCTCAACTCCAA 500
189 GGCACAAACTCAAAAGTGGGGAGCAAGTTCGGCTGAGAGCTGAGCTTCTCAACTCAGA 248
501 CTTGCAAGCTCTCAAGAGGAGCTGAGAGAGCTCAGCGGTGGGCTGAGACCTTGGCGGCA 560
249 CTTGCAAGTCTCAAGAGAGCTGAGAGGCTGAGACATCTCGTGGGCTCTATCAGAA 308
561 TGGAGGGAAGCTGTCACTGTCCCATGATCCCGCTGGGCTGAGAGAGACCAAGAGAGCT 620
309 CACAGAGGAGGCAATACGATTCCTGATTCCTTGGGCTGAAAGGAAACGAAAGAGCT 368
621 GGAATGCTACACCGCTGAAGAGCTGATCTCAGTGCACCTTTGGAGAGACGCGGCTC 680
369 CGACTTTGCAAGTCTCAAGGATTTATCTTGGAAACATTAAGTGAAGATGGCTATTT 428
681 CTAAGAGGCAAGAAATCAGGAGCTGAGAGGCTTGGCGAGGCCATGCGGACCCCGAGCCG 740
429 ATATGAAGATGAATTCAGATCTTATGATCTGAGACAGCTTGTTCGAGCGCTAGCGG 488
741 GAATGAGTGGGCTGAGAGCTGTCAAGCGTATTAAACAGCTGTGCTTCTGATGC 800
489 GGATGAGGCGGGGTGAACTGTGTATGATGACATATTCATCAGCTGGGCTTTGTTCGAG 548
801 GCGCTTCTCAGCCCTGCCAGAGGCTTCGGGCTTCTTCCACTGTGACACTGCGTTAC 860
549 TCGATTTCTCCGGCCACAGCGAGATGGGACTCTCTGTTCACCTGGTATGACTCTCTAC 608
861 TGGGGTCCCGCCAGCAGCGTGGCTTGGGCTTTCGAGAAGGCGAGCGTCTCTTCAACAT 920
609 CGGGTTCGCTCAGCCAGCAGAACCTGCTGAGAGAGCGGAGTCTCTGTTTCAACAC 668
921 CGGTGCTTCCACAGCAGATTTGGGCGCGCAGGACCGCTCTGACCGAGGTTGCCG 980
669 TGGGGCTCTTACACCCAGATTTGGGACCGGCTGTGATCGGACAGCGAGCTGGGCTGGA 728
981 CCGGCTATGAGGCTTTCAGAGGCGGCTGGGCGCTTCAGCGCTCTGAGGAGAGAACTT 1040
729 GAGTGCCATGATGCTTTTTCAGAGAGCGCAGGGGTTTAAATACCTGAAAGACATTT 788
1041 CTTCCATGCGCGGAGCCAGACATGAGCGTGTGCTGCTTGGCGCATGAGAGAGCTCAT 1100
789 TACCATACTCCAAATTCAGACATGAGCCCTGCCATGCTCAGCGTCTCGTCAAAATGAT 848
1101 GATGGCCCGCCAGGAAATGTGTGTTGAGGCGCTCTCACCACCTGCTCCATGCGCCC 1160
849 GCTTGCAACAGCCCAAGAAACGCTTTTGAGAAATCAGGCTTCTCTG-----GGATCCG 902
1161 CCAAGACTGTGTCGCCAGCTGCGCTGGGCGCAGAGCGCGCCAGGTGGCAGCGCAGTA 1220
903 GAATGAATTTCTCATGCTGTGTAAGTGGCTCAGAGAGGCTGCTAAGTGGAGAGGCTA 962
1221 CAGGCTAGTGCACGGGACCATGCGCCAGCCACCCGTCACAGCTACGTGCTGCTCTCTG 1280
963 CCAACAGCTACACGACGATGAGCCAGGCGCGGTGAAAGAGAAATCCCTACTCTCTG 1022
1281 GACTGCCCTGTTGATGATGAGCGGAGTACTTCCGCTCCCTGCGCCACTACCACTGAGC 1340
1023 GGCAGCTTATGCTGCTGAAGGCGCCACCTACGCGGCTTGGCCACTTCTTCTCTG 1082
1341 CATGGCCCTCTGCAAGCGGCTCCCGAGCGACCGAGGAGAGCTCCCGCAGCAGCAGAGGT 1400
1083 CATCTCTCTCATCGACCAACAGGTGAGCCAGGACCGGATCTGGACCAACAGGAGAGTG 1142
1401 CTTCTCTGAGC-----CCCCCACTCTCTTAAGCCCGAGGCGCTGTGCT--TGCCGAGGA 1454
```



```
QY 321 GGGCTGTGACTCCCTGACGAGATCCAGTGGCGGACAGCTGCGAGCGCGGAGGGCCCGAGAT 380
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
105 GGGCTGTAAATCCCTTTGCAACAAACCGGCGGAGTAAATTCGAGAAATCAAGAGCTGCTTT 164
QY 381 TCACGAGCAGATTGACAAGGAGCTGCAGATGGGAGCGGCGCTGCAACCTCTACAGAGC 440
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
165 GAATCAGCAGATCTCTGAAGCCATGCGGATGAGGACCGGAGCGGAAACCTTCTGAAAGT 224
QY 441 CACCGACAAACACCGGTTGAGAGAGCGGTGCGCCCTGGAGCTGAGCTCAACTCCAA 500
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
225 GGGCCACAAACTCAAGAGTGGGAGCAAGTGGGCTGGAGCTGAGCTTCTGCAACTCAGA 284
QY 501 CTTGAGCTGCTGAGGAGAGCTGAGAGCTCAGCGGTGGGTGAGGACCTTGGCGGCA 560
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
285 CTTGAGATGCTCAAGGAGAGCTGAGGGGCTGAACATCTCGTGGGCGCTTATCAGAA 344
QY 561 TGGGAGCGAAGCTGTCACTGTCCCATGATCCCTTGGCCCTGAAGGAGCAAGAGCT 620
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
345 CACAGAGGAGCATTTACGATTCCTGATTTCTCTTGGCCCTGAAGGAAACGAAGAGCT 404
QY 621 GGACTGTCTACACCGCTGAAGAGCTGATCTAGTGCACTTTTGGAGAGAGCGGCGCTC 680
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
405 CGACTTTCAGTTCGCTCAAGGATTTTATCTGGAACTTACAGTGAAGATGCTATTT 464
QY 681 CTACGAGGAGAAATCAGGAGCTGGAGGCTTGGGAGGCGCATCGGAGCCCCAGCGG 740
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
465 ATATGAAGATGAAATGCAAGATCTTATGGATCTGAGACAAAGCTTGTGCGAGCGCTAGCGG 524
QY 741 GAATGAGTGGGCGCTGGAGCTGCTCAACGCTTATTACAAACAGCTGTGCTTCTGATGC 800
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
525 GGAATGAGGCGGGTGGAACTGCTGATGACATCTTATCAGCTGGGCTTTGTGAGAG 584
QY 801 GGGTTCCTACACCTGCGAGAGGCTCGGGCTCTTTCCACTGGTACGACTCGCTTAC 860
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
585 TCGATTTCTCCGCGCCACACGCGCAGATGCGACTCTCTGTTTCACTGATGACTCTCTAC 644
QY 861 TGGGCTCCGCGCCAGCAGCTGCGCTGCGCTTCGAGAGGCGAGGTTCTCTTCAACT 920
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
645 CGGGTTCGGGTACGCCAGCAGAACTGCTGTGGAGAGGCGAGTGTCTCTTCAACAC 704
QY 921 CGTGCCTTCCACACGAGATTTGGGCGCGCGCAGACCGCTCTCTCACCGAGGTTGCCG 980
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
705 TGGGCGCTCTACACCCAGATTGGACCGGCTGTGATCGGAGCGAGGCTGGGCTGGA 764
QY 981 CGCGCTATGAGGCGCTTCCAGAGGCGCGCTGGGCGCTTCAGCCCTCTGAGGAGAACTT 1040
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
765 GAGTCCCATAGATGCTTTACAGAGAGCGCGAGGGGTTTTAAATTACCTGAAGACACATT 824
QY 1041 CTCCCATGCGGAGCCAGACATGAGCGCTGCTGCTCTGCGCACTGGAGCAGCTCAT 1100
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
825 TACCCATATCTCAAGTTACGACATGAGCCCTGCCATGCTCAGCGTGTCTCGTCAAAATGAT 884
QY 1101 GATGCGCCAGGCGCCAGGAATGTGTTTGGGGCTCTCACCACTGCTCTCATCGCCCC 1160
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
895 GCTTGCAAGGCCAAGAAAGGTGTTGAGAAATCAGCTTCTCTG-----GATCGG 938
QY 1161 CCAAGACTGCTGGCCCAAGCTGCGCTGGCGAGAGCGCCGCCAGGTGGCAGCGAGTA 1220
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
939 GAATGAATTTCTTCATGCTGTGTGAAGTGGCTCAGGAGGCTGTAAAGTGGGAGAGTCTA 998
QY 1221 CAGCTAGTGACCGGACCATGCGCCACCGACCGCTCCAGCACTAGCTGCTGCTCTCTG 1280
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
999 CCAACAGCTACACGAGCGCATGAGCCAGCGCGCGGTGAAGAGAAACATCCCTACTCTG 1058
QY 1281 GACTGCCCTGTGTGATGTCAGGCGGAGTACTTCCGCTCCCTGGCCCACTACCGTAGC 1340
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1059 GGCAGCTTAGCTCGTGAAGGCCACCATACGCGGCCCTTGGCCCACTACTTCACTGC 1118
QY 1341 CATGCGCTCTGCGAGCTCTCCCAAGCAGCGAGGAGAGCTTCCCAAGCAGCAGCAGCT 1400
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1119 CATCTCTCTATCGACACCAAGGTGAAGCCAGGCGAGCTGAGACCAAGGAGAGT 1178
QY 1401 CTTCTGTGAGC-----CCCCCACTCTCTAAGCCCGGAGGCGCTGTGC--TGCGCAGGA 1454
```

```
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1179 CCTGTCCCAGCTCTACGACCATGATCCAGAGGGGCTGACACCTTGGCCACACTGAGAA 1238
QY 1455 GCTGGAGGAGCGCAGCAGCTTGGCAAGCACAACCTGAAGCGTGCCATCTCTGGGGCAGGA 1514
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1239 TGATCAGCAGCGCCGACACAGCTGGGGAAGTCCCACTTTGCGCAGAGCCATGCTCATCAGA 1298
QY 1515 GGAGGCGCTGCGGCTGACGCGCTGTGCGCGCTCTCTCGCGAGGTGGACTGCTTCTGGGC 1574
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1299 GGAGTGGTGGCGGAGGCGAGCTCTGCAAGAGCTGCGGAGCATTGAGGTGCTACAGAA 1358
QY 1575 TGTGATCTCCCAAGACGCTGACGCGTCACTGCGCCAAAGTATGCGGAGCTCGACCGTGAGGA 1634
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1359 GGTGTGTGTGCGGACAGGAAACGCTCCCGGCTCAGCTACGCCCCAGCACCAGAGGAGGA 1418
QY 1635 TGACTTCTGTGAGGCTGCGGAGGCGCGGACATCAGCTTAAGACCCACCAGAGCCAGA 1694
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1419 TGACCTGTGAACCTGATGCGAGCGCCCCAGTGTGTGCTAAACCTGAGCAAGAGGTGA 1478
QY 1695 GGCCAGGATGCCACGCTGTCCAGGGGAAGGGGCTGACATCTTCCATCGGCTGGGGCC 1754
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1479 CATTATATGCCCCAGTTCTCAAGCTGACAGTCAAGAGCTTCTTCCAGAGCTGGGCC 1538
QY 1755 CTTGTCTGTGTTCTCAGCAAGAACCGGTGGCGCTGTGGGGCGCGCTCACCTGACCCG 1814
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1539 CTTATCTGTGTTTTCGGCTAACAGCGGTGGAGCGCTCTCGAAGCATCCGCTTCACTGC 1598
QY 1815 AGGAGAGGCGGCTTTGGCTCACGCTTGGGGAGACTGCGCTGTCTCATCGCTGCCGT 1874
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1599 AGAAGAGGGGACTTTGGGGTTCACTTGAGAGGGAACGCCCCGTTCAAGTTTCACTTCT 1658
QY 1875 CATTCAGAGGAGCCAGGCGCGCGCTGGCTGAAGAGGCGGCACTACATTGTGTCAGT 1934
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1659 GGATCCTTACTGCTGCTGCTGGTGGCAGAGGCCGCGGAGAGAGATTAATTGTCTCAT 1718
QY 1935 GAATGGGAGCAGCATCAGAGTGTGGAGACACGCGAGGTGGTGAAGAGCTGAAGGCTGC 1994
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1719 TCAGCTTGTGATTGTAAGTGGCTGACGCTGAGTGAGGTATGAAGCTGCTGAAGAGCTT 1778
QY 1995 GGGAGAGGCGGGCGCCAGCTGACGCTGAGGTGGTGTGCTGCTGCTG 2034
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1779 TGGCGAGGACGAGATCGAGATGAAGTCTGAGCGCTCCTG 1818
```

RESULT 15
AK098246

LOCUS Homo sapiens cDNA FLJ40927 fis, clone UTERU2006593, weakly similar to GTP-RHO BINDING PROTEIN 1.

ACCESSION AK098246

VERSION AK098246.1 GI:21759222

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K., Arita, M., Musashino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Iehi, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.

NEDO human cDNA sequencing project
Unpublished

2 (bases 1 to 3500)

Isogai, T. and Yamamoto, J.

REFERENCE

Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT

NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB.

FEATURES

source
Location/Qualifiers
1. .3500
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UTERU2006593"
/tissue_type="uterus"
/clone_lib="UTERU2"
/note="cloning vector: pME18SFL3"

ORIGIN

Query Match 20.0%; Score 494.8; DB 9; Length 3500;
Best Local Similarity 56.7%; Pred. No. 9.5e-72;
Matches 976; Conservative 0; Mismatches 732; Indels 12; Gaps 3;
QY 321 GGGCTGTGACTCCCTGACGAGATCCAGTGGCGCCAGCTGCAGAGCGCGAGGGCCAGAT 380
DB 105 GGGCTGTATCCCTTGCACAAACCGCCGGAGTAATTGCAGAAATCAAGAGCTGCTT 164
QY 381 TCACAGCAGATTGACAAAGAGCTGCAGATGCGGAGCGGGCGTGAGAACCTTACAGAGC 440
DB 165 GAATCAGCAGATCTCGAAGCCATGCGGATGAGGACCGGAGCGGAAACCTTCTGAAAGT 224
QY 441 CACAGCAACAAACCGGCTGAGAGAGCGTCCGCTGGAGTGAAGTTCGCTGAGTTCGCT 500
DB 225 GGCACAAAATCAAGGTGCGGAGCAAGTGGCGCTGGAGTGAAGTTCGCTGAGTTCGCT 284
QY 501 CTTGCACTGCTGAAGGAGAGCTGGAGAGAGCTCAGCGGTGGCGTGAGACCTTGGCGGCA 560
DB 285 CTTGCACTGCTCAAGGAGAGCTGGAGGAGCTGAAACATCTCGGTGGCGCTTATCAGAA 344
QY 561 TGGAGCGAAGCTGTCTGCTGCCATGATCCCTCGGCTGAGAGAGCAGAGAGCT 620
DB 345 CACAGAGGAGGCAATTAAGATTCCTGATTCCTCTTGGCTGAAGGAAACGAAAGACGT 404
QY 621 GGAATGCTTACACCGCTGAGGAGCTGATCTCAGTGCACCTTTGGAGAGACGCGCTC 680
DB 405 CGACTTTGCACTGCTCTCAAGGATTTATCTGGAACATTAAGTGAAGTGTCTATT 464
QY 681 CTACAGGCGAAGATCAAGGAGCTGGAGGCGCTCGGCGAGGCCATCGGACCCCGAGCGG 740
DB 465 ATATGAAGATGAATTCAGATCTTATGATCTGAGACAAAGCTTGTGGAGCGCTAGCG 524
QY 741 GAATGAGTGGGCGCTGGAGCTGTCTACAGCTTATACACAGCTGTGCTTCTGGATGC 800
DB 525 GGATGAGCGGGGTGGAATCTGTGATGATGATCTTATCACTGAGCTTGTGCGAGAG 584
QY 801 GCGCTTCTCACCCCTGCGAGGAGCTCGGCTCTTCTTCACTGATGAGCTCGCTTAC 860
DB 585 TCGATTTCTCCGCGCCACACGCGAGATGGAGCTCTGTTACCTGGTATGATCTCTCAC 644
QY 861 TGGGAGTCCCGCCAGCAGCGTGCCTGGCTTCGAGAGGCGAGCGTCTCTTCAACAT 920
DB 645 CGGGTTCGCTCACAGCAGATGGGGCGCGCAGAGCGCTCTGCAACGAGGCTGCGCG 980
QY 921 CGGTGCTCTCACAGCAGATGGGGCGCGCAGAGCGCTCTGCAACGAGGCTGCGCG 980
DB 705 TGGGGCCCTTACACCCAGATGGGACCCGCTGTGATCGGAGAGCGAGGCTGGGCTGGA 764
QY 981 CCGGCTATGAGGCGCTTCCAGAGGCGCTGGGCGCTTCAAGCTCTCTGAGGAGAACTT 1040
DB 765 GAGTGGCATAGATGCCTTTCAGAGAGCGCGAGGGGTTTAAATTTACCTGAAAGACAT 824
QY 1041 CTCGATGCGCGGAGCCAGACATGAGGCGTGGCTGCTGCGCTGCGGCTGAGGAGCTCAT 1100

DB 825 TACCATACTCCAAAGTTACGACATGAGCCCTGCCATGCTCAGCGTGTCTCGTCAAAATGAT 884
QY 1101 GATGGCCAGGCGCCAGGATGTGTGTTGAGGCGCTCTCACACCTGCTCCATGCGCCC 1160
DB 885 GCTTGACAGGCGCAAGAAAGCGTGTGAGAAATCAGCCTTCTCTG-----GGATCCG 938
QY 1161 CCAAGACTGCTGCGCCAGCTGCGCTGCGCGAGGAGCGCCAGGTGCGAGCCCGAGTA 1220
DB 939 GAATGAATTTCTCATGCTGCTGAAGTGGCTCAGGAGGCTGCTAAGGTGGAGAGTCTA 998
QY 1221 CAGGCTAGTGACCGGACCAATGCGCCAGCACCCTGTCAGACTAGTGTGCTGTCTCTG 1280
DB 999 CCAACAGCTACACGACGATGAGCCAGGCGCGGTGAAGAGAACATCCCTTACTCTCTG 1058
QY 1281 GACTGCGCTGCTGATGTCAGAGCGAGTACTTTCGCTCCCTGCGCCACTTACACGTAGC 1340
DB 1059 GGCAGCTTACGCTGAGGCGCCACCTAGCGGCGCTGGCCCTGCTTCTTCTACTGCT 1118
QY 1341 CATGGCCCTCTGCGAGGCTCCCGCAGCAGCCAGGAGAGCTCCCGACGACGAGCAGGT 1400
DB 1119 CATCTCTCTCATGACCAACAGGTGAAGCCAGGACGATCTGGAACCCAGGAGAGTG 1178
QY 1401 CTTCTCTGAGC-----CCCCCACTCTCTTAAGCCCCGAGGCCCTGTGTC--TGCGGAGGA 1454
DB 1179 CTTGTCCCACTCTACGACCAATGCGAGGCGGCTGACACCTTGGCCACACTTACAGAA 1238
QY 1455 GCTGGAGGAGCGAGGAGCTTGGCAAGGCACACCTGAAGCGTGCCTTCTGGGGCAGGA 1514
DB 1239 TGATCAGAGCGCGCAGCTGGGGAAGTCCCACTTGGCAGAGCCATGCTCATCAGGA 1298
QY 1515 GAGGCGCTGCGGCTGCAACGCTGTGCCGCTCTGCGCGAGGTGAGACTGTCTGGGGC 1574
DB 1299 GGAAGTGGTGGGAGCGAGCTCTGCAAGAGTGGCGAGCATTTGAGGTGCTTACAGAA 1358
QY 1575 TGTGATCTCCAGACGCTGAGCGCTCATGCGCCAGTATGCGGAGCTCCAGCTGAGGA 1634
DB 1359 GGTGCTGTGTCGCGCAGAGAACGCTCCCGCTCACTGAGCCAGCAGGAGGAGGA 1418
QY 1635 TGACTTCTGTGAGCTGCGGAGCGCCGAGCATCCAGCTTAAGACCCACAGAGCCAGGA 1694
DB 1419 TGACTGCTGAACCTGATCGACGCGCCAGTGTGTGCTTAAACTGAGCAAGAGTTGA 1478
QY 1695 GGCAGGATGCCACGCTGTCCAGGGAAGGGCGCTGACATCTTCCATCGGCTGGGGCC 1754
DB 1479 CATTTATTTGCCCGAGTTCTCAAGCTGACAGTCAAGGACTTCTTCCAGAGCTGGGCCC 1538
QY 1755 CTTGTCTGTGTTCTCAGCAAGAACCGTGGCGCTGGTGGGCGCTGCCACCTGACCCG 1814
DB 1539 CTTATCTGTGTTTGGCTAAACAGCGGTGGAGCGCTTCTCGAAGCATCCGCTTCACTGC 1598
QY 1815 AGGAGGCGCGCTTTGGCTCAGCTTCCGGGAGAGCTCGCTGTCTCTCATCGCTGCGGT 1874
DB 1599 AGAAGAGGAGACTTGGGGTTTCACTTGAAGGGAACGCCCCGTTTCAAGGTTCATCTCT 1658
QY 1875 CATTCAGGAGCAGCGCGCGCTGGCTGAGGAGGCGGCTTACATTGTGTGCTGCT 1934
DB 1659 GGAATCTTACTGCTCTGCTCGTGGCAGGAGCGCGGGAAGAGATATATTGTCTCCAT 1718
QY 1935 GAATGGGCGCCTACGAGTGTGGAGACACGCGGAGGTGTGACGAGCTGAAGGCTGCG 1994
DB 1719 TCAGCTTGTGATTTGAAGTGGCTGACGCTGAGTGAAGTTATGAAGCTGCTGAAGAGCTT 1778
QY 1995 GGAAGAGCGCGCGCCAGCTTCAAGCTGCGAGGTGTGTGCTGCTGCTG 2034
DB 1779 TGGCGAGGAGAGATCGAGATGAAGTCGTGAGCGCTCTCTG 1818

Search completed: July 13, 2004, 06:09:55
Job time : 6352.83 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2004, 20:25:00 ; Search time 4177.82 Seconds

(without alignments)

17647.909 Million cell updates/sec

Title: US-10-697-266-1

Perfect score: 2469

Sequence: 1 tgcggccaggtgtgctgcgg.....aaaaaaaaaaaaaaaaaaaaa 2469

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba.*
2: em_esthum.*
3: em_estmu.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_est3.*
12: gb_est4.*
13: gb_est5.*
14: gb_estfun.*
15: em_estom.*
16: em_gss_hum.*
17: em_gss_inv.*
18: em_gss_pln.*
19: em_gss_vrt.*
20: em_gss_fun.*
21: em_gss_mam.*
22: em_gss_mus.*
23: em_gss_pro.*
24: em_gss_rod.*
25: em_gss_phg.*
26: em_gss_vrl.*
27: gb_gss1.*
28: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	903.4	36.6	1070	13 BX400896	BX400896 BX400896
C 2	850	34.4	1017	13 BX444630	BX444630 BX444630
C 3	849.4	34.4	1201	13 BX447754	BX447754 BX447754
C 4	812.6	32.9	1201	9 AL578313	AL578313 AL578313

5	802.4	32.5	1001	13	BQ072325
6	796.4	32.3	1135	11	BC012135
7	787.6	31.9	899	13	BU171128
8	742	30.1	775	14	CA432432
9	722	29.2	1127	13	EX341691
10	716.6	29.0	1201	13	EX375208
11	715.8	29.0	1201	13	EX375208
12	665.2	26.9	703	14	CB851722
13	665	26.9	707	14	CB850862
14	626.2	25.4	650	13	BU629871
15	624.4	25.3	1201	13	EX384556
16	613.2	24.8	1243	12	BM928275
17	605.6	24.5	679	10	BE251330
18	588.8	23.8	607	14	CD513476
19	556	22.5	1067	12	BG339711
20	553.8	22.4	612	9	AI917311
21	544.4	22.0	3058	11	AK004849
22	541.2	21.9	3351	11	AK050214
23	533.6	21.6	830	12	BI518788
24	526	21.3	986	12	BI770906
25	519.2	21.0	611	10	BE328691
26	515.6	20.9	1109	12	BM548005
27	513.2	20.8	793	12	BI103148
28	508.6	20.6	542	10	AW874410
29	506	20.5	1201	9	AL555948
30	502.4	20.3	523	10	AW295415
31	502.4	20.3	926	13	BQ880416
32	477.8	19.4	493	12	BM193145
33	473.4	19.2	494	12	BM153139
34	468.6	19.0	709	10	BE408252
35	467.8	18.9	609	14	CB216237
36	466.6	18.9	506	9	AI624807
37	461	18.7	755	13	BQ179567
38	458.6	18.6	810	10	BE607519
39	446	18.1	1408	12	BM475275
40	437.6	17.7	490	10	BF348343
41	435.4	17.6	489	9	AA292035
42	426	17.3	1021	13	BX444631
43	423.2	17.1	721	13	BQ442874
44	411	16.6	955	13	BQ927119
45	403	16.3	999	12	EM802482

ALIGNMENTS

RESULT 1
BX400896/c
LOCUS BX400896 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION CDNA clone CS0DK004YE05 3-PRIME, mRNA sequence.
ACCESSION BX400896
VERSION BX400896.1 GI:30630225
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1070)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4751.f For more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DK004AC03NP1&cluster=4751.f. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600


```
Db 818 AGCGCTCACTGGCAAGTATGCGAGCTCGACCGTGAGGATGACTTCTGTGAGGCTGCCG 759
QY 1655 AGCCCCGACATCCAGCCTAAGACCCACAGAGCCAGAGGCCAGATGCCACGCTGT 1714
Db 758 AGCCCCGACATCCAGCCTAAGACCCACAGAGCCAGAGGCCAGATGCCACGCTGT 699
QY 1715 CCCAGGGGAAGGGGCTGACATCTTCATCCATCGGCTGGGGCCCTGTGTGTCTCAGCCA 1774
Db 698 CCCAGGGGAAGGGGCTGACATCTTCATCCATCGGCTGGGGCCCTGTGTGTCTCAGCCA 639
QY 1775 AGAACCGGTGGCGGCTGTGTGGGCGCGTCCACTGACCCAGAGAGGGGCGCTTTGGCC 1834
Db 638 AGAACCGGTGGCGGCTGTGTGGGCGCGTCCACTGACCCAGAGAGGGGCGCTTTGGCC 579
QY 1835 TCACGCTTCGGGGAGACTCGCTGTCTCATCGCTGCGCTCATCTTCAGGGAGCCAGGCGG 1894
Db 578 TCACGCTTCGGGGAGACTCGCTGTCTCATCGCTGCGCTCATCTTCAGGGAGCCAGGCGG 519
QY 1895 CGGCGGCTGGCTGAAGAGGGGCACTAATTTGTGTCAATGATGAGGAGCCATGAGGT 1954
Db 518 CGGCGGCTGGCTGAAGAGGGGCACTAATTTGTGTCAATGATGAGGAGCCATGAGGT 459
QY 1955 GGTGAGAGACACCGGAGGTGTGACGAGCTGAAGGCTCGAGGCTCGGAGAGGGGCGGCGG 2014
Db 458 GGTGAGAGACACCGGAGGTGTGACGAGCTGAAGGCTCGGAGAGGGGCGGCGGCGG 399
QY 2015 TGCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2074
Db 398 TGCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 339
QY 2075 TCCTGCTGGGCCCCAGGGGCTTCTAAGAGCCAGAGGAGCATGTTGTGAAGACCCCGG 2134
Db 338 TCCTGCTGGGCCCCAGGGGCTTCTAAGAGCCAGAGGAGCATGTTGTGAAGACCCCGG 279
QY 2135 CATCCAGCTGGGCCCCAGGGGCTTCTAAGAGCCAGAGGAGCATGTTGTGAAGACCCCGG 2194
Db 278 CATCCAGCTGGGCCCCAGGGGCTTCTAAGAGCCAGAGGAGCATGTTGTGAAGACCCCGG 219
QY 2195 AGACTGGAGGTGCCCCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2254
Db 218 AGACTGGAGGTGCCCCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 160
QY 2255 AGACCCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2314
Db 159 AGACCCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 101
QY 2315 GGCAGCAAGCAGGAGCATGCCCTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2374
Db 100 GGCAGCAAGCAGGAGCATGCCCTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 42
QY 2375 CCTCATGCTG 2384
Db 41 CCTCATGCTG 32
```

```
RESULT 3
BX447754/c
LOCUS
DEFINITION BX447754 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
ACCESSION BX447754
VERSION BX447754.1 GI:31035860
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
```

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4751.f For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0AH001CB05NP1&cluster=4751.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS0AH001CB05NP1.

FEATURES

source

1..1201
/location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DH001YD09"
/tissue_type="T CELLS (JURKAT CELL LINE)"
/cell_line="JURKAT CELL LINE"
/note="Vector: Homo sapiens T CELLS (JURKAT CELL LINE)"
/note="Vector: PCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match	34.4%;	Score 849.4;	DB 13;	Length 1201;
Best Local Similarity	95.6%;	Pred. No. 5.9e-90;		
Matches 936;	Conservative 8;	Mismatches 24;	Indels 11;	Gaps 7;

QY	1459	GAGGAGCGCAGGACGCTTGGCAAGCGCACACCTGAAAGCGTGCCATCTCTGGGGCAGGAGG	1518
Db	976	GAAGAGCGCAGGACGCTTGGCAAG--MACACCTTAGCGTGCCATCTCTGGGGCAGGAGG	919
QY	1519	GGCTGCGGCTGACGCCCC--TGTCGGCGCTCTGCGGAGGTGACCTGCTGCGGCTGT	1577
Db	918	GGCTGCGGCTGACGCCCCCTTGTGCGCGCTCTGCGGAGGTGACCTGCTGCGGCTGT	859
QY	1578	GATCTCCAGAGCGCTGACGCGCTCACTGCCCAAGTATCGGAGCTCGACCTGAGGATGA	1637
Db	858	GATCTCCAGAG--GCTGCGAGCGCTCACTGGCCAAAGTATCGGAGCTCGACCGTGAAGATGA	800
QY	1638	CTTCTGTGAGGCTGCGGAGGCCCCGGACATCAGCCTTAAGACCCACAGAGCCAGAGGC	1697
Db	799	CTTCTGTGAGGCTGCGGAGGCCCCGGACATCAGCCTTAAGACCCACAGAGCCAGAGGC	740
QY	1698	CAGATGCCAGCTGTCTCCAGGGGAAAGGGGCTGACATCTTCCATCGGCTGGGCCCCCT	1757
Db	739	CAGATGCCAGCTGTCTCCAGGGGAAAGGGGCTGACATCTTCCATCGGCTGGGCCCCCT	680
QY	1758	GTCTGTGTTCTCAGCCAGAACCGGTGGCGGCTGGTGGGGCCCGTCCACCTGACCCGAG-	1816
Db	679	GTCTGTGTTCTCAGCCAGAACCGGTGGCGGCTGGTGGGGCCCGTCCACCTGACCCGAG	620
QY	1817	GAGAGGGCGCTTTGGCCTCAGCCTTTCGGGGAGACTGCGCTGTCTCTATCCGCTGCGGTCA	1876
Db	619	GAGAGGGCGCTTTGGCCTTACGCTTCGGGGAGACTGCGCTGTCTCTATCCGCTGCGGTCA	560
QY	1877	TTCAGGAGGAGCAGGCGCGGCTGCTGCTGAAGAGGGGCGCACTAATTGTGACGTGA	1936
Db	559	TTCAGGAGGAGCAGGCGCGGCTGCTGCTGAAGAGGGGCGCACTAATTGTGACGTGA	500
QY	1937	ATGGGAGCCATGACGTGTGTGGAGACACCGGAGGAGGTGGTACGAGCTGAGGCTGCGG	1996
Db	499	ATGGGAGCCATGACGTGTGTGGAGACACCGGAGGAGGTGGTACGAGCTGAGGCTGCGG	440
QY	1997	GAGAGGGGGCGCCAGGCTGTCAGGTGTGTGCTGTGCTGCCAGCTCTAGACTGCCAGCT	2056
Db	439	GAGAGGGGGCGCCAGGCTGTCAGGTGTGTGCTGTGCTGCCAGCTCTAGACTGCCAGCT	380
QY	2057	TGGGGAGCCGCGGCGCTGCTGTGCTGCGGCCCGCCAGGGGCTTCTAAGAGCCAGAGGAGC	2116


```

VERSION      BQ072325.1  GI:19901371
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgabbs-remail.nih.gov
              Tissue Procurement: Life Technologies, Inc.
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM12810 row: b column: 21
              High quality sequence stop: 623.
FEATURES     Location/Qualifiers
              1..1001
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:5761460"
               /lab_host="DH10B"
               /clone_lib="NIH MGC 122"
               /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
               Site 1: NotI; Site 2: EcoRV (destroyed); RNA source
               anonymous pool of 24 week female lung, 16 week female
               spleen, and 20-22 week male spleens. Library is oligo-dT
               primed and directionally cloned (EcoRV site is destroyed
               upon cloning). Average insert size 1.4 kb, insert size
               range 1-3 kb. Library is normalized and enriched for
               full-length clones and was constructed by C. Gruber
               (Invitrogen). Research Genetics tracking code 026. Note:
               this is a NIH_MGC Library."
ORIGIN
Query Match      32.5%; Score 802.4; DB 13; Length 1001;
Best Local Similarity 96.4%; Pred. No. 1.9e-84;
Matches 864; Conservative 0; Mismatches 26; Indels 6; Gaps 4;
QY 321 GGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGGCAGCTGCAGAGCCGCGGCCAGAT 380
DB 42 GGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGGCAGCTGCAGAGCCGCGGCCAGAT 101
QY 381 TCACCAGCAGATTGACAGGAGCTGCAGATCGGACGGCGCTGAGAACCTCTACAGAGC 440
DB 102 TCACCAGCAGATTGACAGGAGCTGCAGATCGGACGGCGCTGAGAACCTCTACAGAGC 161
QY 441 CACCAGCAACACCGGGTGGAGAGACGCTGCCTGCGAGCTGAGCTACGTCACCTCCAA 500
DB 162 CACCAGCAACACCGGGTGGAGAGACGCTGCCTGCGAGCTGAGCTACGTCACCTCCAA 221
QY 501 CCTGAGCTGTGAAGGAGGAGCTGGAGAGCTCAGCGGTGGCGTGGACCTTGGCGCGCA 560
DB 222 CCTGAGCTGTGAAGGAGGAGCTGGAGAGCTCAGCGGTGGCGTGGACCTTGGCGCGCA 281
QY 561 TGGGAGCGAAGCTGTCACTGTCCCATGATCCCCCTGGGCTGAGAGAGACAAGAGAGCT 620
DB 282 TGGGAGCGAAGCTGTCACTGTCCCATGATCCCCCTGGGCTGAGAGAGACAAGAGAGCT 341
QY 621 GGACTGGTCTACCGCTGAGGAGCTGATCTCAGTGACATTTGGAGAGAGGGGCTC 680
DB 342 GGACTGGTCTACCGCTGAGGAGCTGATCTCAGTGACATTTGGAGAGAGGGGCTC 401
QY 681 CTACGAGCAGAAATCAGGAGCTGGAGGCGCTCGCGGAGCGCATGCGGACCCCGCGG 740
DB 402 CTACGAGCAGAAATCAGGAGCTGGAGGCGCTCGCGGAGCGCATGCGGACCCCGCGG 461

```

741 GAATGAGTCGGGCTGGAGCTGCTCACAGCTATTACAAACAGGCTGTGCTTCTCGATGC 800
462 GAATGAGTCGGGCTGGAGCTGCTCACAGCTATTACAAACAGGCTGTGCTTCTCGATGC 521
801 GGGCTTCTTCCACCCCTGCCAGAGGCTCGGGCTCTTTTCCACTGTGTAAGACTCGCTTAC 860
522 GGGCTTCTTCCACCCCTGCCAGAGGCTCGGGCTCTTTTCCACTGTGTAAGACTCGCTTAC 581
861 TGGGCTTCCCGGCCCAGAGCGTGGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 920
582 TGGGCTTCCCGGCCCAGAGCGTGGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 641
921 CGGTGCTTCCACACGCGAGATTGGGCGCGCCAGAGACCGCTCTCTGACCGAGGAGGAGGAG 980
642 CGGTGCTTCCACACGCGAGATTGGGCGCGCCAGAGACCGCTCTCTGACCGAGGAGGAGGAG 701
981 CGGTGCTTCCACACGCGAGATTGGGCGCGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1039
702 CGGTGCTTCCACACGCGAGATTGGGCGCGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 761
1040 TCTCCATGCGCG-AGCCAGACATGAGCGTGGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAG 1098
762 TCTCCATGCGCG-AGCCAGACATGAGCGTGGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAG 821
1099 ATGATGGCGCCAGGCGCCAGAGATTGTGTGTTG--AGGGGCTCTTCCACACCTGCTCCATGG 1156
822 ATGATGGCGCCAGGCGCCAGAGATTGTGTGTTG--AGGGGCTCTTCCACACCTGCTCCATGG 881
1157 CCCCCCAGAGCTGCTGG--CCAGCTGCGCTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1210
882 GCGGCGGAGCTGGCTGGGCGCCAACTGGCGGCTGGGCGGAGGAGGAGGAGGAGGAGGAGGAG 937

BC012135 1135 bp mRNA linear HTC 19-NOV-2003
Homo sapiens rhophilin, Rho GTPase binding protein 1, mRNA (cdna
clone IMAGE:4555079), with apparent retained intron.
BC012135
BC012135.1 GI:15082433
HTC.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1135)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Snailus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
PUBMED 12477932
2 (bases 1 to 1135)
Strausberg, R.
Direct Submission
Submitted (02-AUG-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ruben Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fiehl, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 29 Row: d Column: 17
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 19882236
This clone has the following problem: retained intron.

FEATURES

Location/Qualifiers
1. .1135
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:455079"
/tissue_type="Uterus, leiomyosarcoma"
/clone_lib="NIH_MGC_46"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

ORIGIN

Query Match 32.3%; Score 796.4; DB 11; Length 1135;
Best Local Similarity 99.9%; Pred. No. 9e-84;
Matches 797; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 321 GGGCTGTGACTCCTGTACGAGAGCTCCAGTGGCGGACGTGCGAGCGCGAGGCCAGAT 380
DB 162 GGGCTGTGACTCCTGTACGAGAGCTCCAGTGGCGGACGTGCGAGCGCGAGGCCAGAT 221
QY 381 TCACACGAGATTGACAGGAGCTGCAGATGCGGACGGGCGCTGAGAACCTCTACAGAGC 440
DB 222 TCACACGAGATTGACAGGAGCTGCAGATGCGGACGGGCGCTGAGAACCTCTACAGAGC 281
QY 441 CACAGCAACACCGGTGAGAGAGCGTGGCCCTGGAGTGTAGCTCACTCACTCAA 500
DB 282 CACAGCAACACCGGTGAGAGAGCGTGGCCCTGGAGTGTAGCTCACTCACTCAA 341
QY 501 CTTGACGCTGTGAAGGAGAGCTGGAGGAGCTCAGCGGTGGCGTGGACCTGCGCCGCGCA 560
DB 342 CTTGACGCTGTGAAGGAGAGCTGGAGGAGCTCAGCGGTGGCGTGGACCTGCGCCGCGCA 401
QY 561 TGGAGCGAAGCTGTCTCTGTCCCATGATCCCTCTGGGCTGAAGGAGACCAAGGAGCT 620
DB 402 TGGAGCGAAGCTGTCTCTGTCCCATGATCCCTCTGGGCTGAAGGAGACCAAGGAGCT 461
QY 621 GGACTGTCTACCGCTGAGGAGCTGTACTAGTGCACCTTTTGGAGGAGCGCGGCTC 680
DB 462 GGACTGTCTACCGCTGAGGAGCTGTACTAGTGCACCTTTTGGAGGAGCGCGGCTC 521
QY 681 CTACGAGCGAGAAATCAGGAGCTGGAGGCGCTCGGCGAGGCCCATCGGACCCCGAGCG 740
DB 522 CTACGAGCGAGAAATCAGGAGCTGGAGGCGCTCGGCGAGGCCCATCGGACCCCGAGCG 581
QY 741 GAATGAGTCGGGCTGGAGCTGTCTACAGGCTATATACACGAGCTGTCTCTGGATGC 800

DB 582 GAATGAGTCGGGCTGGAGCTGTCTACAGGCTATTAAACACGAGCTGTCTCTGGATGC 641
QY 801 GGGCTTCTCTACCCCTGCCAGAGCGCTCGGGCTCTTCTTCCACTGTGTAGACTCGCTTAC 860
DB 642 GGGCTTCTCTACCCCTGCCAGAGCGCTCGGGCTCTTCTTCCACTGTGTAGACTCGCTTAC 701
QY 861 TGGGGTCCCGGCGCCAGCAGCGTGGCTTTCGAGAAAGGGCAGCGTCTTCTTCAACAT 920
DB 702 TGGGGTCCCGGCGCCAGCAGCGTGGCTTTCGAGAAAGGGCAGCGTCTTCTTCAACAT 761
QY 921 CGGTGCGCTCCACACGAGATTGGGGGCGGCCAGGACCGGTCTCTGCACCGAGGCTGCCG 980
DB 762 CGGTGCGCTCCACACGAGATTGGGGGCGGCCAGGACCGGTCTCTGCACCGAGGCTGCCG 821
QY 981 CGCGCTATGGAGGCGCTTCCAGAGGCGCGTGGGGCTTTCAGGCTCTCTGAGGCGAGACTT 1040
DB 822 CGCGCTATGGAGGCGCTTCCAGAGGCGCGTGGGGCTTTCAGGCTCTCTGAGGCGAGACTT 881
QY 1041 CTCCCATGGCGCGGAGCCCGACATGAGCGTGGCTTCTTGGCACTGGAGCGACTCAT 1100
DB 882 CTCCCATGGCGCGGAGCCCGACATGAGCGTGGCTTCTTGGCACTGGAGCGACTCAT 941
QY 1101 GATGGCGCGGCGCCAGGA 1118
DB 942 GATGGCGCGGCGCCAGGA 959
RESULT 7
BUL71128
LOCUS
DEFINITION BUL71128 899 bp mRNA linear EST 04-SEP-2002
AGENCOURT_7935660 NIH_MGC_67 Homo sapiens cdna clone IMAGE:6146655
5', mRNA sequence.
ACCESSION BUL71128
VERSION BUL71128.1 GI:22685112
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 899)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM13475 Row: d Column: 16
High quality sequence stop: 563.
FEATURES
Location/Qualifiers
source 1. .899
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6146655"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site: 1; NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Library constructed by Life Technologies."

ORIGIN

Query Match 31.9%; Score 787.6; DB 13; Length 899;
Best Local Similarity 98.6%; Pred. No. 1e-82;
Matches 826; Conservative 0; Mismatches 9; Indels 3; Gaps 3;


```
QY 1125 GTTTCAGGCGCTCTCACACAGCTGCTCCATGCGCCGCCAAGACTGCTGTGGCCACGCTGGG 1184
Db |||||
QY 1185 CTTGCGCGAGGAGGCGCGCCAGGTGGAGCCGAGTACAGGCTAGTGACCGGACCATGGC 1244
Db |||||
QY 61 CTTGCGCGAGGAGGCGCGCCAGGTGGAGCCGAGTACAGGCTAGTGACCGGACCATGGC 120
QY 1245 CGAGCACCCGTCACGAGTACGTGCTGTCTTCCTGGACTGCGCTGTGGTCAATGTCAGAGC 1304
Db |||||
QY 1305 CGAGTACTTCCGCTCCTCGGCCACTACACATGACCATGCGCTTGGCGGCTCGCC 1364
Db |||||
QY 1365 AGCGACCGAGGAGAGCTCCCCACGACGAGCAGGTCTTCTGACGCGCCGCCACCTCTCTC 1424
Db |||||
QY 241 AGCGACCGAGGAGAGCTCCCCACGACGAGCAGGTCTTCTGAGCGCCGCCACCTCTCTC 300
QY 1425 TAAGCCCGAGGCGCTGTGTCGCGCAGAGCTGAGAGGCGCAGGAGCTTGGCAAGGC 1484
Db |||||
QY 301 TAAGCCCGAGGCGCTGTGTCGCGCAGAGCTGAGAGGCGCAGGAGCTTGGCAAGGC 360
QY 1485 ACACCTGAAGCGTGCCATCTCGGCGCAGGAGGCGCTGCGGCTGCGAGCGCTTGGCG 1544
Db |||||
QY 361 ACACCTGAAGCGTGCCATCTCGGCGCAGGAGGCGCTGCGGCTGCGAGCGCTTGGCG 420
QY 1545 CGTCTGCGGAGGTGGACCTGCTTTCGCGGCTGTGATCTCCAGAGCGTGCAGCGCTCACT 1604
Db |||||
QY 421 CGTCTGCGGAGGTGGACCTGCTTTCGCGGCTGTGATCTCCAGAGCGTGCAGCGCTCACT 480
QY 1605 GCGCAAGTATGCGAGCTGCGAGTGCAGCTGAGGATGACTTCTGTGAGGCTGCGAGGCGCCGGA 1664
Db |||||
QY 481 GCGCAAGTATGCGAGCTGCGAGCTGAGGATGACTTCTGTGAGGCTGCGAGGCGCCGGA 540
QY 1665 CATCAGCGCTAAGACCCACAGAGCCAGAGCCAGGATGCCAGCGCTGTCAGGCGAA 1724
Db |||||
QY 541 CATCAGCGCTAAGACCCACAGAGCCAGAGCCAGGATGCCAGCGCTGTCAGGCGAA 600
QY 1725 GGGG-CCTGACATCTTCCATCGGCTGGGCGCCCTGTCTGTGTCTCAGCCAAAGACCGGT 1783
Db |||||
QY 601 GGGGCGCTGACATCTTCCATCGGCTGGGCGCCCTGTCTGTGTCTCAGCCAAAGACCGGT 660
QY 1784 GCGGCGCTGGTGGGCGCCCTGTCACCTGACCCGAGGAGAGGCGGCTTGGCTTCAAGCTTC 1843
Db |||||
QY 661 GCGGCGCTGGTGGGCGCCCTGTCACCTGACCCGAGAGAGGCGGCTTGGCTTCAAGCTTC 720
QY 1844 GGGGAGACTCGGCTGTCTCTCATGCTGCGGCTCATTTCCAGGAGGAGCCAGGCGCGG-CGGCT 1902
Db |||||
QY 721 GGGGAGACTCGGCTGTCTCTCATGCTGCGGCTCATTTCCAGGAGGAGCCAGGCGCGGCTT 780
QY 1903 GGCCTGAAGGAGGCGGACTACATT-GTGTAGTGAATGGGCGAGCCATGACAGTGTGTGG 1959
Db |||||
QY 781 GGCCTGAAGGAGGCGGACTACCTTGGGCTGAGTGAATGGGCGAGCCATGCGAGGCGGGG 838

RESULT 8
CA432432/c
LOCUS
DEFINITION
  UI-H-FL1-bgu-e-15-0-UI-s1 NCI CGAP_Fl1 Homo sapiens cDNA clone
  UI-H-FL1-bgu-e-15-0-UI 3', mRNA sequence.
ACCESSION
  CA432432
VERSION
  1
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 775)
AUTHORS
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
```

JOURNAL
COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: James Martin
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes

FEATURES
source

Location/Qualifiers
1..775
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FL1-bgu-e-15-0-UI"
/tissue_type="Cell lines"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP_Fl1"
/note="Organ: Chondrosarcoma; Vector: pT7T3-Pac
(Pharmacia) with a modified polylinker; Site 1: EcoR I;
Site 2: Not I; NCI CGAP_Fl1 is a normalized cDNA library
derived from a pool of mRNA obtained from 4 cell lines
from grade III chondrosarcoma tissues. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
GAGTCTGGTG. The cell lines were provided by Dr. James
Martin from the University of Iowa.
TAG_TISSUE=Human Chondrosarcoma Grade 3 cell line mix
TAG_LIB=UI-H-FL1
TAG_SEQ=GAGGTCGGTG"

ORIGIN

Query Match 30.1%; Score 742; DB 14; Length 775;
Best Local Similarity 99.6%; Pred. No. 2.2e-77;
Matches 775; Conservative 0; Mismatches 0; Indels 3; Gaps 3;
QY 1677 GACCCACAGAGGCGCAGAGGATGCCAGCTGTGCCAGGAGGCGGCTGACAT 1736
Db |||||
QY 1737 CTTCCATCGGCTGGGCGCCCTGTCTGTGTTCTCAGCCAAAGACCGGTGGCGGCTGGTGGG 1796
Db |||||
QY 715 CTTCCATCGGCT-GGGCGCCCTGTCTGT-TTCTCAGCCAGAACCGGTGGCGGCTGGTGGG 658
QY 1797 GCGCGTCCACTGACCGAGAGAGGCGGCTTTGGGCTCAGCTTCGGGGAGACTCGCC 1856
Db |||||
QY 657 GCGCGTCCACTGACCGGAGAGAGGCGGCTTTGGGCTCAGCTTCGGGGAGACTCGCC 598
QY 1857 TGTCTCATCGCTCGCTCATTTCCAGGAGGCGCGCGCGCTGGCTTGAAGGAGG 1916
Db |||||
QY 597 TGTCTCATCGCTCGCTCATTTCCAGGAGGCGCGCGCGCTGGCTTGAAGGAGG 538
QY 1917 CGACTACATTGTGTGAGTGAATGGGCGAGCCATGACAGTGTGTGGAGACACGCGGAGGTGGT 1976
Db |||||
QY 537 CGACTACATTGTGTGAGTGAATGGGCGAGCCATGACAGTGTGTGGAGACACGCGGAGGTGGT 478
QY 1977 GACGAGCTGAAGGCTCGGGAGAGGCGGCGCAGCTTCAGGTGTGTGCTGTGCTGCC 2036
Db |||||
QY 477 GACGAGCTGAAGGCTCGGGAGAGGCGGCGCAGCTTCAGGTGTGTGCTGTGCTGCC 418
QY 2037 CAGCTCTAGACTGCCAGCTTGGGGAGACCGCGCGCGCTTCTGCTGGGCGCCACAGGGGCT 2096
Db |||||
QY 417 CAGCTCTAGACTGCCAGCTTGGGGAGACCGCGCGCGCTTCTGCTGGGCGCCACAGGGGCT 358

```
QY 2097 TCTAAGAGCCAGAGGAGCATGGTTGCAAGACCCCGGCATCCAGTGGGCGCAGTCCCCG 2156
Db 357 TCTAAGGAGCCAGAGGAGCATGGTTGCAAGACCCCGGCATCCAGTGGGCGCAGTCCCCG 298
QY 2157 GCCCCTCCCTCAACTGAGCGCGAAGGCCAGAGCCAGAGCTGGAGGCTGCCCGCCAGCC 2216
Db 297 GCCCCTCCCTCAACTGAGCGCGAAGGCCAGAGCCAGAGCTGGAGGCTGCCCGCCAGCC 238
QY 2217 CTGTGCCCCAGTGAAGCCAGCTCCGGCCCTCATCTTTGAAGCACCCAGGGTGGCGGTGAGG 2276
Db 237 CTGTGCCCCAGTGAAGCCAGCTCCGGCCCTCATCTTTGAAGCACCCAGGGTGGCGGTGAGG 178
QY 2277 GCCAGATCCCTGACGCCCTGAGCCTGCTGCTCCAGCTGCGCAGCAGCACCGAGCATGCC 2336
Db 177 GCCAGATCCCTGACG - CCTCAGCCCTGGCTCCAGCTGCGCAGCAGCACCGAGCATGCC 119
QY 2337 CTCCCCACCCAGAGGACCTCCGGGCAATGCTGTCCGCCCTCATGCTGGAGGCTGCCCTCG 2396
Db 118 CTCCCCACCCAGAGGACCTCCGGGCAATGCTGTCCGCCCTCATGCTGGAGGCTGCCCTCG 59
QY 2397 GGCACCTGCTGCCATTAAGACTGTGTAGACTGTCTGAAAAAATAAAAAA 2454
Db 58 GGCACCTGCTGCCATTAAGACTGTGTAGACTGTCTGAAAAAATAAAAAA 1
```

```
RESULT 9
BX341691 1127 bp mRNA linear EST 01-MAY-2003
LOCUS BX341691 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DK004YE05 5-PRIME, mRNA sequence.
ACCESSION BX341691
VERSION BX341691
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1127)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4751.f For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DK004AC03QF1
cluster=4751.f Contact : feng liang Email : fliang@lifetech.com
URL : http://fulllength.invitrogen.com/InvitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DK004AC03QF1.
```

```
FEATURES
source
1..1127
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK004YE05"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

```
Query Match 29.2%; Score 722; DB 13; Length 1127;
Best Local Similarity 95.5%; Pred. No. 4.2e-75;
Matches 806; Conservative 10; Mismatches 20; Indels 8; Gaps 7;
QY 321 GGGCTGTGACTCCCTGACGCAGATCCAGTGGCGCCAGCTGCAGAGCCGAGGCCAGAT 380
```

```
Db 267 GGGCTGTGACTCCCTGACGCAGATCCAGTGGCGCCAGCTGCAGAGCCGCA -GGCCCCAGAT 325
QY 381 TCACACGACGATGACAAAGGAGCTGCAGATGCGGACGGGGCTGAGAACTCTTACAGAC 440
Db 326 TCACACGACGATGACAAAGGAGCTGCAGATGCGGACGGGGCTGAGAACTCTTACAGAC 385
QY 441 CACCACAAACACCGGGTGCAGAGAGCGTTCGCCCTGGAGCTGAGCTACGCTCAACTCCAA 500
Db 386 CACCACAAACACCGGGTGCAGAGAGCGTTCGCCCTGGAGCTGAGCTACGCTCAACTCCAA 445
QY 501 CCTGACAGCTGCTGAAGGAGAGCTGCAGAGAGCTTCAGCGGTGGCGTGGACCTGGCCGCA 560
Db 446 CCTGACAGCTGCTGAAGGAGAGCTGCAGAGAGCTTCAGCGGTGGCGTGGACCTGGCCGCA 505
QY 561 TGGGAGGAGAGCTGTCTCATGTGCCCATGATCCCGCTGGGCTGAAGGAGACCAAGAGCT 620
Db 506 TGGGAGGAGAA -CTGTCACTGTCCCATGATCCCGCTGGGCTGAAGGAGACCAAGAGCT 564
QY 621 GGACTGCTTACACCCCTGAAGGAGCTGATCTCAGTGCACCTTTGGAGAGAGCGCGCTC 680
Db 565 GGACTGCTTACACCCCTGAAGGAGCTGATCTCAGTGCACCTTTGGAGAGAGCGCGCTC 624
QY 681 CTACGAGGACAGAAATCAGGAGAGCTGAGGCGCTTGGCGAGGCCATCGGGAACCCAGCCG 740
Db 625 CTACGAGGACAGAAATCAGGAGAGCTGAGGCGCTTGGCGAGGCCATCGGGAACCCAGCCG 684
QY 741 GAATGAGTGGGCGCTCGAGCTGCTCACAGCTATTACACACAGCTGTGCTTCTTGATGC 800
Db 685 GAATGAGTGGGCGCTCGAGCTGCTCACAGCTATTACACACAGCTGTGCTTCTTGATGC 744
QY 801 GCGCTTCTTCCACCCCTGCCAGAGAGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC 860
Db 745 GCGCTTCTTCCACCCCTGCCAGAGAGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC 804
QY 861 TGGGGTCCCGGCCAGAGCGTGCCTGGCTTTCGAGAGAGGAGCGAGCTTCTTTCACAT 920
Db 805 TGGGGT -CGGGCCAGCAGCGTGCCTTGGCTTTCGAGAGAGGAGCGAGCTTCTTTCACAT 863
QY 921 CGGTGCGCTTCCACACGACAGATTCGGGCGCGCCAGGACCGCTCTCGACACGAGGCTGCCG 980
Db 864 CGGTGCGCTTCCACACGACAGATTCGGGCGCG -CAGGACCGCTCTCTGACCCAGAGGT -SCCG 921
QY 981 CCGCGCTATGGAGGCGCTTCCAGAGGCGCGCTGGGGCGCTTCAGCTCTGAGGGAGAACTT 1040
Db 922 CCGCGCTATGGAGG -CTTCCAGAGGCGCGCTGGGGCGCTTCAGCTCTGAGGGAGAACTT 980
QY 1041 CTCCTATGCGCGCCAGCGCCAGACATGAGCGCTGCGTCCCTCTGGCACTGGAGCGAGCTCAT 1100
Db 981 CTCCTATGCGCGCCAGCGCCAGACATGAGCGCTGCTCTCTGCGCACTGGGAGCGAGCTCAT 1040
QY 1101 GATGGCCCGAGCCCGAGAAATGCTGTTTGGAGGCGCTCTCACACCTGCTCCATGGGCCCC 1160
Db 1041 GATGGCCCGAGCCCGAGAAATGCTGTTTGGAGGCGCTCTCACACCTGCTCCATGGGCCCC 1098
QY 1161 CCAA 1164
Db 1099 CAAA 1102
```

RESULT 10

```
BX375208 1201 bp mRNA linear EST 08-MAY-2003
LOCUS BX375208 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DC010YG11 5-PRIME, mRNA sequence.
ACCESSION BX375208
VERSION BX375208
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1201)
```

```

AUTHORS      Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE        Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      Contact: Genoscope
              Genoscope - Centre National de Sequencage
              BP 191 91006 EVRY cedex - France
              Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
              Library was constructed by Life Technologies, a division of
              Invitrogen. This sequence belongs to sequence cluster 4751.f For
              more information about this cluster, see
              http://www.genoscope.cns.fr/
              cgi-bin/cluster.cgi?seq=CS0DH001AD06QPl&cluster=4751.f. Contact :
              Feng Liang Email : fliang@lifetech.com URL :
              http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
              Faraday Avenue Genoscope sequence ID : CS0DH001AD06QPl.
FEATURES     Location/Qualifiers
              1..1201
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="CS0DH001YD09"
              /tissue_type="T CELLS (JURKAT CELL LINE)"
              /cell_line="JURKAT CELL LINE"
              /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
              /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
              with a NotI-oligo (dT) primer. Five prime end enriched,
              double-strand cDNA was digested with Not I and cloned into
              the Not I and EcoRV sites of the pCMVSPORT 6 vector.
              Library was not normalized."
ORIGIN
source
Query Match      29.0%; Score 715.8; DB 13; Length 1201;
Best Local Similarity 91.9%; Pred. No. 2.1e-74;
Matches 824; Conservative 6; Mismatches 13; Indels 54; Gaps 5;

QY 921 CGGTGCTCCACACGACGAGA-TTGGGGCGCGCCAGACCGCTCTCTGCACCGAGGTGCC 979
    |||||
Db 822 CGGTGCTCCACACGACGAGA-TTGGGGCGCGCGCCAGACCGCTCTCTGCACCGAGGTGCC 881
    |||||

QY 980 GCGGCGCTATGAGAGCGCTTCAGAGGGCGCGCTGGGGCGCTTCAGCCCTCTTGAGGGAGAACT 1039
    |||||
Db 882 GCGGCGCTATGAGAGCGCTTCAGAGGGCGCGCTGGGGCGCTTCARCTCTCTTGAGGGAGA--C 939
    |||||

QY 1040 TCTCCCATGCGCGGAGCCAGACATGAGCGCTGCGTCCCTCTGCGCACTGGAGCAGCTCA 1099
    |||||
Db 940 TTYTCCATGCTSCGAGCCARACATGAGCGCTS---STCCCTCTGCCACTGGAGAGCTCA 996
    |||||

QY 1100 TGATGGCCAGGCGCCAGGAATGTGTGTTGAGGGCGCTCTCACC 1142
    |||||
Db 997 TGATGGCCW--GGCCARGAATGTGTTTTRAGGCTTTCACMC 1037
    |||||

RESULT 11
BX405203      1201 bp      mRNA      linear      EST 13-MAY-2003
BX405203      Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
clone CS0DH001YD09 5-PRIME, mRNA sequence.
ACCESSION    BX405203
VERSION      BX405203.1 GI:30635313
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 1201)
AUTHORS      Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE        Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      Contact: Genoscope
              Genoscope - Centre National de Sequencage
              BP 191 91006 EVRY cedex - France
              Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
              Library was constructed by Life Technologies, a division of
              Invitrogen. This sequence belongs to sequence cluster 4751.f For
              more information about this cluster, see
              http://www.genoscope.cns.fr/
              cgi-bin/cluster.cgi?seq=CS0AH001CB05QPl&cluster=4751.f. Contact :
              Feng Liang Email : fliang@lifetech.com URL :
              http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
              Faraday Avenue Genoscope sequence ID : CS0AH001CB05QPl.
FEATURES     Location/Qualifiers
              1..1201
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="CS0DH001YD09"
              /tissue_type="T CELLS (JURKAT CELL LINE)"
              /cell_line="JURKAT CELL LINE"
              /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
              /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
              with a NotI-oligo (dT) primer. Five prime end enriched,
              double-strand cDNA was digested with Not I and cloned into
              the Not I and EcoRV sites of the pCMVSPORT 6 vector.
              Library was not normalized."
ORIGIN
Query Match      29.0%; Score 715.8; DB 13; Length 1201;
Best Local Similarity 91.9%; Pred. No. 2.1e-74;
Matches 824; Conservative 6; Mismatches 13; Indels 54; Gaps 5;

QY 329 ACTCGCTGACCGACATCCAGTGGCGCCAGCTGCAGAGCCGCGAGGCCAGATTCCACGAC 388
    |||||
Db 211 ACCCGCGCTCGACATCCAGTGGCGCCAGCTGCAGAGCCGCA-GGCCAGATTCCACGAC 269
    |||||

QY 389 AGATTGACAAGGAGCTGCAGATGCGAGCGCGCTGAGAACTCTACAGAGCCACCGACA 448
    |||||
Db 270 AGATTGACAAGGAGCTGCAGATGCGAGCGCGCTGAGAACTCTACAGAGCCACCGACA 329
    |||||

```


Db 345 CAGAGGGAGCATGGTTGCAAGACCCCGGCATCACTGGGGCAGTCCCCCGGCCCTTCCTC 286
QY 2167 AACTGAGCCGAAAGCCAGCAGGCAAGACTGGAGGTGCCCCCAGAGCCCTGTGCCCA 2226
Db 285 AACTGAGCCGAAAGCCAGCAGGCAAGACTGGAGGTGCCCCCAGAGCCCTGTGCCCA 226
QY 2227 GTGAAGCCAGCTCCGCGCTCATCTTGAAGCACCAGGTTGGCCGTGAGGGCCAGATCC 2286
Db 225 GTGAAGCCAGCTCCGCGCTCATCTTGAAGCACCAGGTTGGCCGTGAGGGCCAGATCC 166
QY 2287 CTGACGCGCTCAGCCCTCGCTCCAGTGGCAGCAAGCAGCATGCCCTCCCAACC 2346
Db 165 CTGACG-CCTCAGCGCTGGCTCCAGCTGGCAGCAAGCAGCATGCCCTCCCAACC 107
QY 2347 AGAGGACCTCCGGCAATGCTGTCCGCTCATGCTGGAGGTGCTCGGCACCTGCTCC 2406
Db 106 AGAGGACCTCCGGCAATGCTGTCCGCTCATGCTGGAGGTGCTCGGCACCTGCTCC 47
QY 2407 TGCCCATTAAGACTGGTCCAGACCTGTCTGAAGCAAGCAAGCAAGCAAGCAAGCA 2452
Db 46 TGCCCATTAAGACTGGTCCAGACCTGTCTGAAGCAAGCAAGCAAGCAAGCAAGCA 1

RESULT 13

CB850862/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

CB850862 707 bp mRNA linear EST 22-APR-2003
UI-CF-EN1-adb-i-17-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
UI-CF-EN1-adb-i-17-0-UI 3', mRNA sequence.

CB850862

CB850862.1

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 707)

Ronald M.F., Lennon G. and Soares M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

97044477

889548

Contact: McCray, PB

McCray Lab

University of Iowa

2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866

Fax: 319 356 7171

Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 FORWARD

POLYA=Yes.

Location/Qualifiers

1..707

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-CF-EN1-adb-i-17-0-UI"

/tissue_type="Primary Lung Cystic Fibrosis Epithelial

Cells"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone_lib="UI-CF-EN1"

/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;

UI-CF-EN1 is a normalized cDNA library containing the

following tissue(s): Primary Lung Cystic Fibrosis

Epithelial Cells. The library was constructed according to

Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CTGCTCAGGT.

TAG_SEQ=None found"

ORIGIN

Query Match 26.9%; Score 665; DB 14; Length 707;
Best Local Similarity 98.7%; Pred. No. 2.1e-68;
Matches 701; Conservative 0; Mismatches 6; Indels 3; Gaps 3;
QY 1750 GGGCCCTCTGCTGTTCTCAGCCCAAGAACCGGTGGCGCTGTCACCTG 1809
Db 707 GGGCCCTCTGCTGTTCTCAGCCCAAGAACCGGTGGCGCTGTCACCTG 650
QY 1810 ACCGAGGAGAGGGCGGCTTTGGCTTCCAGCTTCGGGAGAGCTCGCTCTCTCATCGCT 1869
Db 649 ACCGAGGAGAGGGCGGCTTTGGCTTCCAGCTTCGGGAGAGCTCGCTCTCTCATCGCT 590
QY 1870 GCCGTCAATTCAGGAGCCAGCCCGCGCTGCTGAAAGAGGGGAGCTACATTGTG 1929
Db 589 GCCGTCAATTCAGGAGCCAGCCCGCGCTGCTGAAAGAGGGGAGCTACATTGTG 530
QY 1930 TCAGTGAATGGCAGCAGCATGCAAGTGTGGAGACACAGGAGGTGTGACGAGCTGAAG 1989
Db 529 TCAGTGAATGGCAGCAGCATGCAAGTGTGGAGACACAGGAGGTGTGACGAGCTGAAG 470
QY 1990 GTTCGGGAGAGGGCGGCGCCAGCTGCAGGTGTGTGCTGCTGCCAGCTCTAGACTG 2049
Db 469 GTTCGGGAGAGGGCGGCGCCAGCTGCAGGTGTGTGCTGCTGCCAGCTCTAGACTG 410
QY 2050 CCCAGCTTGGGGAGCCGCGCGCTCTGCTGTGGCCCGCCAGGGGCTTCTAAGGAGCCAG 2109
Db 409 CCCAGCTTGGGGAGCCGCGCGCTCTGCTGTGGCCCGCCAGGGGCTTCTAAGGAGCCAG 350
QY 2110 AGGAGCATGTTTGAAGACCCCGCATCCAGTGGGCGAGTCCCGGCGCTCTCTCAAC 2169
Db 349 AGGAGCATGTTTGAAGACCCCGCATCCAGTGGGCGAGTCCCGGCGCTCTCTCAAC 290
QY 2170 TGGAGCCGAAAGGGCCAGCAGGCGCAAGCTGAGAGCTGCCCCCAGCCCTGTGCCCCAGTG 2229
Db 289 TGGAGCCGAAAGGGCCAGCAGGCGCAAGCTGAGAGCTGCCCCCAGCCCTGTGCCCCAGTG 230
QY 2230 AAGCCAGCTCCGCTCTCATCTTGAAGCACCAGCGGTGGCGGTGAGGGCCAGGATCCCTG 2289
Db 229 AAGCCAGCTCCGCTCTCATCTTGAAGCACCAGCGGTGGCGGTGAGGGCCAGGATCCCTG 170
QY 2290 CACGCCCTCAGCCCTGCTCCAGCTGGCAGCAAGCAGCAGGATGCTCCCTCCCAAGCAGA 2349
Db 169 CACG-CCTCAGCCCTGCTCCAGCTGGCAGCAAGCAGCAGGATGCTCCCTCCCAAGCAGA 111
QY 2350 GGACCTCCGGGCAATGCTGTCCCGCTCATGCTGGAGGTGCTGCTGGGCGAGCTGCCTGC 2409
Db 110 GGACCTCCGGGCAATGCTGTCCCGCTCATGCTGGAGGTGCTGCTGGGCGAGCTGCCTGC 51
QY 2410 CCATTAAAGACTGGTTCAGACTGCTGTGAAAAAAGCAAGCAAGCAAGCAAGCAAGCA 2459
Db 50 CCATTAAAGACTGGTTCAGACTGCTGTGAGCCCAAGCAAGCAAGCAAGCAAGCAAGCA 1

RESULT 14

BU629871/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

BU629871 650 bp mRNA linear EST 23-SEP-2002
UI-H-FL0-bdr-a-08-0-UI.s1 NC1 CGAP FL0 Homo sapiens cDNA clone
UI-H-FL0-bdr-a-08-0-UI 3', mRNA sequence.
BU629871
BU629871.1 GI:23296343
EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 650)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-k@mail.nih.gov
 Tissue Procurement: James Martin
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
 Seq Primer: M13 FORWARD
 POLYA=Yes

FEATURES
 source Location/Qualifiers
 1..650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-FL0-bdr-a-08-0-UI"
 /tissue_type="Cell lines"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP FLO"
 /note="Organ: Chondrosarcoma; Vector: p7T3-Pac (Pharmacia) with a modified polylinker; Site 1; Ecor I; Site 2; Not I; NCI CGAP FLO is a cDNA library derived from a pool of mRNA obtained from 4 cell lines from grade III chondrosarcoma tissues. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally into p7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GAGTCCGCTG. The cell line was provided by Dr. James Martin from University of Iowa.
 TAG_TISSUE=Human Chondrosarcoma Grade 3 cell line mix
 TAG_LIB=UI-H-FL0
 TAG_SEQ=GAGTCCGCTG"

ORIGIN
 Query Match 25.4%; Score 626.2; DB 13; Length 650;
 Best Local Similarity 98.6%; Pred. No. 7.2e-64;
 Matches 642; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
 QY 1809 GACCCGAGGAGAGGCGGCTTTGGCCCTCAGCTTCGGGGAGACTCGGCTGTCTTCATCGC 1868
 Db 650 GACCCGAGGAGAGGCGGCTTTGGCCCTCAGCTTCGGGGAGACTCGGCTGTCTTCATCGC 591
 QY 1869 TGCCGTGATTCAGGAGACCGAGCCGCGGCTGTGGCTGAAGAGGCGGACTACATGT 1928
 Db 590 TGCCTGATTCAGGAGACCGAGCCGCGGCTGTGGCTGAAGAGGCGGACTACATGT 531
 QY 1929 GTCACTGAATGGGAGCATTCAGTGTGTGAGACACGCGAGTGTGTGACGGAGCTGAA 1988
 Db 530 GTCACTGAATGGGAGCATTCAGTGTGTGAGACACGCGAGTGTGTGACGGAGCTGAA 471
 QY 1989 GGCTGCGGAGAGGCGGCGCCAGCTTCAGTGTGTGCTGTGCGCCAGCTTAGACT 2048
 Db 470 GGCTGCGGAGAGGCGGCGCCAGCTTCAGTGTGTGCTGTGCGCCAGCTTAGACT 411
 QY 2049 GCCAGCTTGGGGNACCGCGGCGGCTTCCTGCGGCCCGAGGGGCTTCTAAGAGCCA 2108
 Db 410 GCCAGCTTGGGGNACCGCGGCGGCTTCCTGCGGCCCGAGGGGCTTCTAAGAGCCA 351

QY 2109 GAGGGAGCATGTTGCAAGACCCCGGCATCCAGTGGGGCAGTCCCGGCCCTCTCAA 2168
 Db 350 GAGGGAGCATGTTGCAAGACCCCGGCATCCAGTGGGGCAGTCCCGGCCCTCTCAA 291
 QY 2169 CTGAGCCGAAAGGCGCCAGCAGGGCAAGACTGAGAGGTGCCCCCAGCCCTGTGCCAGT 2228
 Db 290 CTGAGCCGAAAGGCGCCAGCAGGGCAAGACTGAGAGGTGCCCCCAGCCCTGTGCCAGT 231
 QY 2229 GAAGCCAGCTCCCGCTCATCTTGAAGACCCAGAGGTGGCCGTGAGGGCCAGGATCCCT 2288
 Db 230 GAAGCCAGCTCCCGCTCATCTTGAAGACCCAGAGGTGGCCGTGAGGGCCAGGATCCCT 171
 QY 2289 GCAGCCCTCAGCCCTGGCTCCAGCTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2348
 Db 170 GCAGC-CCTCAGCCCTGGCTCCAGCTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 112
 QY 2349 AGGACCTCCCGGCAATGCTGTCCCGCTCATCTTGAAGACCCAGAGGTGGCCGTGAGGGCCAGGATCCCT 2408
 Db 111 AGGACCTCCCGGCAATGCTGTCCCGCTCATCTTGAAGACCCAGAGGTGGCCGTGAGGGCCAGGATCCCT 52
 QY 2409 CCATTAAGACTGGTTCAGACCTGTCTGAAAAAAGAAAAAAGAAAAA 2459
 Db 51 CCATTAAGACTGGTTCAGACCTGTCTGAAAAAAGAAAAAAGAAAAA 1

RESULT 15
 BX384556
 LOCUS
 DEFINITION BX384556 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
 cDNA clone CS0DK011YM10 5-PRIME, mRNA sequence.
 ACCESSION BX384556
 VERSION BX384556.1 GI:30440364
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1201)
 Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DK011BG050P1.

FEATURES
 source Location/Qualifiers
 1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DK011YM10"
 /cell_type="HELA CELLS COT 25-NORMALIZED"
 /cell_line="HELA"
 /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
 /notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and Ecor V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Query Match 25.3%; Score 624.4; DB 13; Length 1201;
 Best Local Similarity 97.3%; Pred. No. 9.5e-64;
 Matches 651; Conservative 4; Mismatches 11; Indels 3; Gaps 2;
 QY 570 AGCTGTCACTGTCCCCATGATCCCCCTGGGCTTGAAGAGACCAAGAGCTGGACTGTC 629
 Db 365 AASTTCACTGTCCCCATGATCCCCCTGGGCTTGAAGAGACCAAGAGCTGGACTGTC 424
 QY 630 TACACCTGAAAGGAGCTGATCTCAGTGCATTTGGAGAGACCGGCGCTCTCTAGAGGC 689


```
Db      425  TACACCGCTGAAGAGCTGATYTCAGTGCACCTTTGGAGAGCGGCGCCTTCCTACGAGGC 484
QY      690  AGAAATCAGGAGCTGGAGGCCCTGGCGCAGGCGCATGCGGACCCCGAGCGGAATGAGTC 749
Db      485  AGAAATCAGGAGCTGGAGGCCCTGGCGCAGGCGCATGCGGACCCCGAGCGGAATGAGTC 544
QY      750  GGGCCTGGAGCTGCTCACAGCCTATTACAAACAGCTGTGCTTCTGGAATGCGCGCTTCCT 809
Db      545  GGGCCTGGAGCTGCTCACAGCCTATTACAAACAGCTGTGCTTCTGGAATGCGCGCTTCCT 604
QY      810  CACCCCTGCCAGGAGCTCGGGCTCTTCTTCCACTGTGAGACTCGCTTACTGGGTCCC 869
Db      605  CACCCCTGCCAGGAGCTCGGGCTCTTCTTCCACTGTGAGACTCGCTTACTGGGTCCC 664
QY      870  GGCCCAAGCAGCGTGCCTTCGCCCTTCGAGAGGCGCAGCGTTCTCTTCAACATCGGTGCCCT 929
Db      665  GGCCCAAGCAGCGTGCCTTCGCCCTTCGAGAGGCGCAGCGTTCTCTTCAACATCGGTGCCCT 724
QY      930  CCACACGAGATTGGGGCGGCGCAGGACCGCTCCTGCACGAGGGTGCCTGCGCGCTAT 989
Db      725  CCACACGAGATTGGGGCGGCGCAGGACCGCTCCTGCACGAGGGTGCCTGCGCGCTAT 784
QY      990  GGAGGCTTCCAGAGGCGCGCTGGGGCTTTCAGCCTCCTGAGGGAGAACTTCTCCCATGC 1049
Db      785  GGAGGCTTCCAGAGGCGCGCTGGGGCTTTCAGCCTCCTGAGGGAGAACTTCTCCCATGC 844
QY      1050  GCCGAGCCGAGACATGAGCGCTGGGTCCCTCTGCGCACTGGAGCAGCTCATGATGGCCCA 1109
Db      845  GCCGAGCCGAGACATGAGCGCTGGGTCCCTCTGCGCACTGGAGCAGCTCATGATGGCCCA 904
QY      1110  GGGCCAGGAATGTGTTTGGGGCTCTCACCACTGCTCATGCGCGCGCGCGCGCGCTG 1169
Db      905  GGGCCAGGAATGTGTTTGGGGCTCTCACCACTGCTCATGCGCGCGCGCGCGCGCTG 963
QY      1170  CCTGGCCAGCTGGCGCTGGCGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTAGT 1229
Db      964  CTTTGGCCAGCTGGCGCTGGCGAGGAGGCG--GCCAGGTGGCAGCGGAGTACAGGCTAGT 1021
QY      1230  GCACGGAC 1238
Db      1022  GCACGGAC 1030
```

Search completed: July 13, 2004, 08:20:59
Job time : 4187.32 secs

his Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 13, 2004, 06:10:50 ; Search time 1005.06 Seconds
(without alignments)

11973.025 Million cell updates/sec

Title: US-10-697-266-1

Perfect score: 2469

Sequence: 1 tcgcggccaggtggtgcgg.....aaaaaaaaaaaaaaaaaaaaa 2469

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3183909 seqs, 2436941669 residues

Total number of hits satisfying chosen parameters: 6367818

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2.*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2469	100.0	2469	13	US-10-697-266-1
2	2469	100.0	2469	13	US-10-274-878-1
3	2105	85.3	2837	13	US-10-451-207-11
4	2082.4	84.3	2738	13	US-10-112-944-74
5	957	38.8	1013	15	US-10-023-896-15
6	955.4	38.7	1013	9	US-09-925-301-174
7	955.4	38.7	1013	15	US-10-023-896-43
8	946.8	38.3	1155	9	US-09-833-361-1293
9	674	27.3	1046	13	US-10-112-944-558
10	496.4	20.1	2100	10	US-10-239-607-52
11	496.4	20.1	3526	16	US-10-239-607-51
12	494.8	20.0	2109	13	US-10-092-900A-351
13	494.8	20.0	3500	16	US-10-108-260A-2397
14	493.2	20.0	2061	9	US-09-895-040A-2

15	493.2	20.0	3484	9	US-09-895-040A-1	Sequence 1, Appli
16	483.8	19.6	3518	10	US-09-814-353-20684	Sequence 20684, A
17	463.6	18.8	2310	13	US-10-092-900A-349	Sequence 349, App
18	418.2	16.9	2757	16	US-10-378-029-49	Sequence 49, Appli
19	368	14.9	19025	13	US-10-697-266-3	Sequence 3, Appli
20	368	14.9	19025	15	US-10-274-878-3	Sequence 3, Appli
21	149.6	6.1	2300	16	US-10-108-260A-469	Sequence 469, App
22	108.4	4.4	2925	16	US-10-094-749-256	Sequence 256, App
23	108.4	4.4	4212	13	US-10-425-114-26843	Sequence 26843, A
24	108.4	4.4	5200	17	US-10-433-794-21	Sequence 21, Appli
25	103.6	4.2	599	10	US-09-764-891-244	Sequence 244, App
26	94.6	3.8	439	10	US-09-918-995-4314	Sequence 4314, Ap
27	93.4	3.8	338	9	US-09-815-343-97	Sequence 97, Appli
28	87	3.5	479	9	US-09-867-701-4120	Sequence 4120, Ap
29	72.4	2.9	500	9	US-09-895-040A-29	Sequence 29, Appli
30	71	2.9	167	16	US-10-239-607-62	Sequence 62, Appli
31	70.8	2.9	501	10	US-09-814-353-15780	Sequence 15780, A
32	70.8	2.9	553	10	US-09-814-353-17840	Sequence 17840, A
33	69.4	2.8	167	9	US-09-895-040A-14	Sequence 14, Appli
34	68.2	2.8	2941	17	US-10-424-599-133011	Sequence 133011, A
35	68	2.8	334	13	US-10-437-963-78091	Sequence 78091, A
36	68	2.8	334	13	US-10-027-632-124788	Sequence 124788, A
37	68	2.8	334	16	US-10-027-632-124789	Sequence 124789, A
38	68	2.8	334	16	US-10-027-632-124788	Sequence 124788, A
39	68	2.8	334	16	US-10-027-632-124788	Sequence 124788, A
40	68	2.8	334	16	US-10-027-632-124788	Sequence 124788, A
41	68	2.8	2469	13	US-10-697-266-1	Sequence 1, Appli
42	68	2.8	2469	15	US-10-274-878-1	Sequence 1, Appli
43	68	2.8	2738	13	US-10-112-944-74	Sequence 74, Appli
44	68	2.8	2837	13	US-10-451-207-11	Sequence 11, Appli
45	66	2.7	401	10	US-09-814-353-5166	Sequence 5166, Ap

ALIGNMENTS

RESULT 1

US-10-697-266-1
; Sequence 1, Application US/10697266
; Publication No. US20040067522A1
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: THEREOF
; FILE REFERENCE: CL001099-CIP-DIV2
; CURRENT APPLICATION NUMBER: US/10/697,266
; PRIOR FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: 10/274,878
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-697-266-1

Query Match 100.0%; Score 2469; DB 13; Length 2469;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2469; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TCGCGGCCAGGTGTTGGTGGCGGCCCTAGCCCGCTCGGAGCGCTCGCGAGCGCGG	60
Db	1	TCGCGGCCAGGTGTTGGTGGCGGCCCTAGCCCGCTCGGAGCGCTCGCGAGCGCGG	60
Qy	61	GCTGCTGACCCCGAGGACCCCCAGCGCGGTGGCGGTGATGATCTCTGAGGAGG	120
Db	61	GCTGCTGACCCCGAGGACCCCCAGCGCGGTGGCGGTGATGATCTCTGAGGAGG	120


```
Db 2281 GATCCCTGCACGCTTACGCTTGGCTCAGCTGGAGCAAGCAGCATGCCCTCC 2340
QY 2341 CCACCCAGAGGACCTCCGGCAATGCTGTCCGCTCATGCTGGAGGCTGCCTCCGGCA 2400
Db 2341 CCACCCAGAGGACCTCCGGCAATGCTGTCCGCTCATGCTGGAGGCTGCCTCCGGCA 2400
QY 2401 CTGCTGCTGCCATTAAGACTGTGACACTGTCTGAAAAAAGAAAAAAGAAAAA 2460
Db 2401 CTGCTGCTGCCATTAAGACTGTGCTCAGACCTGTCTGAAAAAAGAAAAAAGAAAAA 2460
QY 2461 AAAAAAAA 2469
Db 2461 AAAAAAAA 2469

RESULT 2
US-10-274-878-1
; Sequence 1, Application US/10274878
; Publication No. US20030049792A1
; GENERAL INFORMATION:
; APPLICANT: ROSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEROF
; FILE REFERENCE: CL001039-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,878
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Human
US-10-274-878-1

Query Match 100.0%; Score 2469; DB 15; Length 2469;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2469; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGCGCCAGGTGTGTGGGGCGGCTAGCCGGCTGCGAGCGCTGCGAGCGGCG 60
Db 1 TCGCGCCAGGTGTGTGGGGCGGCTAGCCGGCTGCGAGCGCTGCGAGCGGCG 60

QY 61 GCTGGCTGACCCCGAGGAGACCCCGAGCGAGCGGTGGCGATGATCCTGGAGAGG 120
Db 61 GCTGGCTGACCCCGAGGAGACCCCGAGCGAGCGGTGGCGATGATCCTGGAGAGG 120

QY 121 CCGAGCGCGCGCGCGCGAGGAGACCCCGCGGCTGCAGATATCTAGAGAAAAA 180
Db 121 CCGAGCGCGCGCGCGCGAGGAGACCCCGCGGCTGCAGATATCTAGAGAAAAA 180

QY 181 AGGAAAAACCTGTGAGCTCTTTACGGGAGAGACCGGAGGCTGAGAGAGCTGTGCG 240
Db 181 AGGAAAAACCTGTGAGCTCTTTACGGGAGAGACCGGAGGCTGAGAGAGCTGTGCG 240

QY 241 TGGAGAGGCTGTGGGTCCACAGAGGGAGAGACCCAGTGTGTCAGCTTGGCCCCATG 300
Db 241 TGGAGAGGCTGTGGGTCCACAGAGGGAGAGACCCAGTGTGTCAGCTTGGCCCCATG 300

QY 301 AATCCGAGCTTTCATGAGTGGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGCAGCTG 360
Db 301 AATCCGAGCTTTCATGAGTGGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGCAGCTG 360

QY 361 CAGAGCCGAGGCGCCAGATTCAACAGCAGATTGACAGGAGCTGCAGATCGGACGGGC 420
Db 361 CAGAGCCGAGGCGCCAGATTCAACAGCAGATTGACAGGAGCTGCAGATCGGACGGGC 420

QY 421 GCTGAGAACCTCTACAGAGCCACACAGCAACACCCGGGTGAGAGAGCGGTGCCCTGGAG 480
```

```
Db 421 GCTGAGAACCTCTACAGAGCCACACCGGGTGAGAGAGCGGTGCCCTGGAG 480
QY 481 CTGAGCTAGCTCAACTCCAACTGAGCTGCTGAAGAGGAGCTGAGAGGAGCTCAGCGGT 540
Db 481 CTGAGCTAGCTCAACTCCAACTGAGCTGCTGAAGAGGAGCTGAGAGGAGCTCAGCGGT 540
QY 541 GCGGTGGAACCTTGGCCGGCATGGAGGAAAGCTGTCACTGTCCCATGATGATCCCTGGGC 600
Db 541 GCGGTGGAACCTTGGCCGGCATGGAGGAAAGCTGTCACTGTCCCATGATGATCCCTGGGC 600
QY 601 CTGAAGGAGACCAAGGAGCTGGAAGCTGTCTACACCGCTGAAGAGCTGATCTCAGTGCAC 660
Db 601 CTGAAGGAGACCAAGGAGCTGGAAGCTGTCTACACCGCTGAAGAGCTGATCTCAGTGCAC 660
QY 661 TTTGGAGAGCGCGGCTCTACGAGGAGGAGAAATCAGGAGCTGAGAGGCTTGGAGGCTTGGGGCAG 720
Db 661 TTTGGAGAGCGCGGCTCTCTACGAGGAGGAGAAATCAGGAGCTGAGAGGCTTGGAGGCTTGGGGCAG 720
QY 721 GCCATGCGGACCCCGAGCGGAAATGAGTCGGGCTTGGAGCTGCTCACAGCTTATTACAAAC 780
Db 721 GCCATGCGGACCCCGAGCGGAAATGAGTCGGGCTTGGAGCTGCTCACAGCTTATTACAAAC 780
QY 781 CAGCTGTGCTTCTCTGATGCGGCTTCTCAACCCCTGCGAGGAGCTTCTTCTTC 840
Db 781 CAGCTGTGCTTCTCTGATGCGGCTTCTCAACCCCTGCGAGGAGCTTCTTCTTC 840
QY 841 CACTGTAGGACTCGCTTACTGGGCTCCCGGCGGAGCAGCTGCTTGGCTTGGAGAG 900
Db 841 CACTGTAGGACTCGCTTACTGGGCTCCCGGCGGAGCAGCTGCTTGGCTTGGAGAG 900
QY 901 GGCAGCGTTCTCTTCAACATCGGTGCTCCACACGAGATTGGGCGCGCGCAGAGCCGC 960
Db 901 GGCAGCGTTCTCTTCAACATCGGTGCTCCACACGAGATTGGGCGCGCGCAGAGCCGC 960
QY 961 TCTTGACCGAGGCTGCGCGCGCTATGAGGCTTCCAGAGGCGCTGCGGCGCTTC 1020
Db 961 TCTTGACCGAGGCTGCGCGCGCTATGAGGCTTCCAGAGGCGCTGCGGCGCTTC 1020
QY 1021 AGCCTCTGAGGAGAACTTCTCCATGCGCGGAGCCAGACATGAGGCTGCGTCCCTC 1080
Db 1021 AGCCTCTGAGGAGAACTTCTCCATGCGCGGAGCCAGACATGAGGCTGCGTCCCTC 1080
QY 1081 TCGCACCTGAGAGCACTCATGATGCGCCCGAGGAAATGTGTGTTTGAAGGCTCTCA 1140
Db 1081 TCGCACCTGAGAGCACTCATGATGCGCCCGAGGAAATGTGTGTTTGAAGGCTCTCA 1140
QY 1141 CCACTGCTCTCATGCGCCCGCAAGACTGCTGCGCCAGCTGCGCTGCGAGAGGCTC 1200
Db 1141 CCACTGCTCTCATGCGCCCGCAAGACTGCTGCGCCAGCTGCGCTGCGAGAGGCTC 1200
QY 1201 GCCAGAGTGGAGCGGAGTACAGGCTAGTGACCGGAGCAGTGGCCAGGACCCGCTCAC 1260
Db 1201 GCCAGAGTGGAGCGGAGTACAGGCTAGTGACCGGAGCAGTGGCCAGGACCCGCTCAC 1260
QY 1261 GACTACGCTGCTGTCTCTGAGCTGCTTGTGATGTCAAGGCGAGTACTTCCGCTCC 1320
Db 1261 GACTACGCTGCTGTCTCTGAGCTGCTTGTGATGTCAAGGCGAGTACTTCCGCTCC 1320
QY 1321 CTGGCCCACTAACAGCTAGCCATGGCCCTCTGCGACGGCTCCCCAGCGAGGAGAG 1380
Db 1321 CTGGCCCACTAACAGCTAGCCATGGCCCTCTGCGACGGCTCCCCAGCGAGGAGAG 1380
QY 1381 CTCCCCAGCAGCAGGAGCTTCTGAGCCCCCGACCTCTTAAGCCCCGAGGCGCT 1440
Db 1381 CTCCCCAGCAGCAGGAGCTTCTGAGCCCCCGACCTCTTAAGCCCCGAGGCGCT 1440
QY 1441 GTGCTGCGCAGGAGCTTGGAGGAGCGCAGGAGCTTGGCAAGGCACACCTGGAAGCGTGC 1500
Db 1441 GTGCTGCGCAGGAGCTTGGAGGAGCGCAGGAGCTTGGCAAGGCACACCTGGAAGCGTGC 1500
QY 1501 ATCTGGGGCAGGAGGAGGCTGCGCTGCAAGCCCTGTGCGCGCTCTGCGGAGGTG 1560
Db 1501 ATCTGGGGCAGGAGGAGGCTGCGCTGCAAGCCCTGTGCGCGCTCTGCGGAGGTG 1560
```

Db 1501 ATCTGGGACGAGGAGGGCTGGGCTGCACGCCCTGTGCGCGCTCTCTGGCGAGGTG 1560
Qy 1561 GACCTCTTGGGCTGTGATCTCCAGACGCTGACGGCTCACTGGCCAAGTATGGGAG 1620
Db 1561 GACCTCTTGGGCTGTGATCTCCAGACGCTGACGGCTCACTGGCCAAGTATGGGAG 1620
Qy 1621 CTCGACGCTGAGATGATCTCTGTGAGGCTGCGAGGCCCGGACATCCAGCCTAAGACC 1680
Db 1621 CTCGACGCTGAGATGATCTCTGTGAGGCTGCGAGGCCCGGACATCCAGCCTAAGACC 1680
Qy 1681 CACAGAGCCAGAGCCAGATGCGACGCTGTCCAGGGGAGGGGCTGACATCTTC 1740
Db 1681 CACAGAGCCAGAGCCAGATGCGACGCTGTCCAGGGGAGGGGCTGACATCTTC 1740
Qy 1741 CATCGGTGGGCCCCCTGTCTGTGTTCTAGCCAAAGAACCGGTGGGCTGGTGGGCC 1800
Db 1741 CATCGGTGGGCCCCCTGTCTGTGTTCTAGCCAAAGAACCGGTGGGCTGGTGGGCC 1800
Qy 1801 GTCCACTGACCGAGAGAGGGCGGCTTTGGCTCAGCGTTCGGGAGACTCGCCTGTC 1860
Db 1801 GTCCACTGACCGAGAGAGGGCGGCTTTGGCTCAGCGTTCGGGAGACTCGCCTGTC 1860
Qy 1861 CTCACTGCTGCCGTCACTCCAGGAGCCAGCGCGCGCTGCGCTGAAAGGGGGCAG 1920
Db 1861 CTCACTGCTGCCGTCACTCCAGGAGCCAGCGCGCGCTGCGCTGAAAGGGGGCAG 1920
Qy 1921 TACATTGTGTCAGTAAATGGGCGCATGAGGTGGGAGACACGCGGAGGTGTCAGC 1980
Db 1921 TACATTGTGTCAGTAAATGGGCGCATGAGGTGGGAGACACGCGGAGGTGTCAGC 1980
Qy 1981 GAGCTGAAGCTGCGGAGAGAGGGCGCGCGCTGAGGTGGTGTGCTGCTGCCCGCAGC 2040
Db 1981 GAGCTGAAGCTGCGGAGAGAGGGCGCGCGCTGAGGTGGTGTGCTGCTGCCCGCAGC 2040
Qy 2041 TCTAGACTGCCAGCTTTGGGGAGCCGCGCGCGCTGCTGCTGGGCCCCAGGGGGCTTCTA 2100
Db 2041 TCTAGACTGCCAGCTTTGGGGAGCCGCGCGCGCTGCTGCTGGGCCCCAGGGGGCTTCTA 2100
Qy 2101 AGGAGCCAGAGGAGGATGTTGTCAGACCCCGGCTCCAGTGGGCGGCTCCCGGCC 2160
Db 2101 AGGAGCCAGAGGAGGATGTTGTCAGACCCCGGCTCCAGTGGGCGGCTCCCGGCC 2160
Qy 2161 CTCCTCAACTGAGCCGAGGAGGAGGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAG 2220
Db 2161 CTCCTCAACTGAGCCGAGGAGGAGGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAG 2220
Qy 2221 GCGGAGTGAAGCCAGCTCCGCGCTCATCTTGAAGCACCCAGGCTGCGCTGAGGAGGCA 2280
Db 2221 GCGGAGTGAAGCCAGCTCCGCGCTCATCTTGAAGCACCCAGGCTGCGCTGAGGAGGCA 2280
Qy 2281 GGATCCCTGCACGCGCTCAGCGCTGCTCCAGCTGGGAGGAGGAGGAGGAGGAGGAGG 2340
Db 2281 GGATCCCTGCACGCGCTCAGCGCTGCTCCAGCTGGGAGGAGGAGGAGGAGGAGGAGG 2340
Qy 2341 CCACCCAGAGGAGCTCCGGGCAATGCTGCTCCGCTCATCTGAGGAGGCTGCTCGGAGCA 2400
Db 2341 CCACCCAGAGGAGCTCCGGGCAATGCTGCTCCGCTCATCTGAGGAGGCTGCTCGGAGCA 2400
Qy 2401 CCTGCTGCCATTAAGATGAGTGGTTCAGACCTGTCTGAAAAAATAAAAAAAAAAAAAA 2460
Db 2401 CCTGCTGCCATTAAGATGAGTGGTTCAGACCTGTCTGAAAAAATAAAAAAAAAAAAAA 2460
Qy 2461 AAAAAAAA 2469
Db 2461 AAAAAAAA 2469

RESULT 3
US-10-451-207-11
; Sequence 11, Application US/10451207
; Publication No. US20040038267A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.

APPLICANT: BURFORD, Neil
APPLICANT: DING, Li
APPLICANT: YUE, Henry
APPLICANT: THORNTON, Michael B.
APPLICANT: CHANLA, Narinder K.
APPLICANT: GANDHI, Ameena R.
APPLICANT: ARVIZU, Chandra S.
APPLICANT: BAUGHN, Mariah R.
APPLICANT: SWARNAKAR, Anita
APPLICANT: DUGGAN, Brendan M.
APPLICANT: LU, Dying Aina M.
APPLICANT: THANGAVELU, Kavitha
APPLICANT: WARREN, Bridget A.
APPLICANT: TANG, Y. Tom
APPLICANT: KHAN, Farrah A.
APPLICANT: YAO, Monique G.
APPLICANT: EMERLING, Brooke M.
TITLE OF INVENTION: INTRACELLULAR SIGNALING MOLECULES
FILE REFERENCE: PF-0868 USN
CURRENT APPLICATION NUMBER: US/10/451,207
CURRENT FILING DATE: 2003-06-18
PRIOR APPLICATION NUMBER: PCT/US01/50315
PRIOR FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: US 60/257,804
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: US 60/260,102
PRIOR FILING DATE: 2001-01-05
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PERL Program
SEQ ID NO 11
LENGTH: 2837
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040038267A1 5560161CB1
US-10-451-207-11

Query Match 85.3%; Score 2105; DB 13; Length 2837;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2116; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 321 GGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGGCGAGCTGAGAGCGCGAGGCGGCGAGAT 380
Db 181 GGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGGCGAGCTGAGAGCGCGAGGCGGCGAGAT 240
Qy 381 TCACCACAGATTTCACAGAGGCTGCGAGATGCGGAGCGGCGCTGAGAACCTTCTACAGAGC 440
Db 241 TCACCACAGATTTCACAGAGGCTGCGAGATGCGGAGCGGCGCTGAGAACCTTCTACAGAGC 300
Qy 441 CACCAGCAACAACCGGCTGAGAGAGCGGCTCGCCCTGGAGCTGAGCTAGCTCAACTCCAA 500
Db 301 CACCAGCAACAACCGGCTGAGAGAGCGGCTCGCCCTGGAGCTGAGCTAGCTCAACTCCAA 360
Qy 501 CTTGACGCTGTGAAGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCT 560
Db 361 CTTGACGCTGTGAAGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCT 420
Qy 561 TGGGAGCGAGCTGTCTGCTGCTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 620
Db 421 TGGGAGCGAGCTGTCTGCTGCTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Qy 621 GGACTGGTCTACACCGCTGAGGAGCTGATCTAGTGCATCTTCTGAGAGGAGCGGCGCTC 680
Db 481 GGACTGGTCTACACCGCTGAGGAGCTGATCTAGTGCATCTTCTGAGAGGAGCGGCGCTC 540
Qy 681 CTACGAGCAGAAATCAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAG 740
Db 541 CTACGAGCAGAAATCAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAG 600
Qy 741 GAATGAGTGGGCGCTTGGAGCTGCTCAGAGCTATTATCAACAGCTGTCTTCTGATGATGATGATGAT 800
Db 601 GAATGAGTGGGCGCTTGGAGCTGCTCAGAGCTATTATCAACAGCTGTCTTCTGATGATGATGATGAT 660

QY 801 GCGCTTCTCAGCCCTGCCAGAGGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC 860
Db |||||
661 GCGCTTCTCAGCCCTGCCAGAGGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC 720
QY 861 TGGGGTCCCGGCCCGCAGCAGCTGCGCCCTGAGAGAGGCGAGGGTCTCTTCAACAT 920
Db |||||
721 TGGGGTCCCGGCCCGCAGCAGCTGCGCCCTGAGAGAGGCGAGGGTCTCTTCAACAT 780
QY 921 CGGTGCTTCCACAGCAGATTGGGGCGCGCCAGAGACCGCTCTCAGCAGAGGGTGGCCG 980
Db |||||
781 CGGTGCTTCCACAGCAGATTGGGGCGCGCCAGAGACCGCTCTCAGCAGAGGGTGGCCG 840
QY 981 CGGTGCTTCCACAGCAGATTGGGGCGCGCCAGAGAGGGCTTTCAGCAGAGGGTGGCCG 1040
Db |||||
841 CGGTGCTTCCACAGCAGATTGGGGCGCGCCAGAGAGGGCTTTCAGCAGAGGGTGGCCG 900
QY 1041 CTCCATGCGCCGAGCCAGACATGAGCGCTGCGTCCCTCTGCGCACTGGAGCAGCTCAT 1100
Db |||||
901 CTCCATGCGCCGAGCCAGACATGAGCGCTGCGTCCCTCTGCGCACTGGAGCAGCTCAT 960
QY 1101 GATGCGCCAGGCCAGAGGATGTTGAGGGCTCTTCAACACCTGCTTCCATGGCCGCC 1160
Db |||||
961 GATGCGCCAGGCCAGAGGATGTTGAGGGCTCTTCAACACCTGCTTCCATGGCCGCC 1020
QY 1161 CCAGACTGCTGCGCCAGCTGCGCTGCGCGCAGAGAGGCGCCAGCTGGAGCGAGTA 1220
Db |||||
1021 CCAGACTGCTGCGCCAGCTGCGCTGCGCGCAGAGAGGCGCCAGCTGGAGCGAGTA 1080
QY 1221 CAGGTAGTGCACCGGACCATGCGCCAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1280
Db |||||
1081 CAGGTAGTGCACCGGACCATGCGCCAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
QY 1281 GACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1340
Db |||||
1341 GACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
QY 1341 CATGCGCCCTGCGCAGCGCTCCAGCAGCAGCGAGGAGCTCCCGCAGCAGAGGCT 1400
Db |||||
1201 CATGCGCCCTGCGCAGCGCTCCAGCAGCAGCGAGGAGCTCCCGCAGCAGAGGCT 1260
QY 1401 CTCTCTGAGCCCGCCAGCTCTTAAGCCCGCGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1460
Db |||||
1261 CTCTCTGAGCCCGCCAGCTCTTAAGCCCGCGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
QY 1461 GGAGCGAGGAGCTTGGCAAGGCAACCTGAAGCGTSCCATCTCTGCGGCGAGGAGGCG 1520
Db |||||
1321 GGAGCGAGGAGCTTGGCAAGGCAACCTGAAGCGTSCCATCTCTGCGGCGAGGAGGCG 1380
QY 1521 GCTCGGCTGCGCAGCGCTGCTGCGGCTGCTGCGGAGTGACCTGCTTGGGCTGTGAT 1580
Db |||||
1381 GCTCGGCTGCGCAGCGCTGCTGCGGCTGCTGCGGAGTGACCTGCTTGGGCTGTGAT 1440
QY 1581 CTCCAGAGCTGCGAGCGCTCACTGGGCAAGTATGCGGAGCTGCGAGCTGCGAGGATGACTT 1640
Db |||||
1441 CTCCAGAGCTGCGAGCGCTCACTGGGCAAGTATGCGGAGCTGCGAGCTGCGAGGATGACTT 1500
QY 1641 CTGTGAGGCTGCGAGGCGCGGAGCTTCAAGCTTCAAGCAGCAGAGGCGAGGCGAG 1700
Db |||||
1501 CTGTGAGGCTGCGAGGCGCGGAGCTTCAAGCTTCAAGCAGCAGAGGCGAGGCGAG 1560
QY 1701 GATGCGAGCTGCTGCGAGGAGGCGCTGAGATCTTCCATGCTGGGCGCCCTGTC 1760
Db |||||
1561 GATGCGAGCTGCTGCGAGGAGGCGCTGAGATCTTCCATGCTGGGCGCCCTGTC 1620
QY 1761 TGTGTTCTCAGCCAGAGAACCGGTGCGGCTGCTGGGCGCCCTGAGCTGCTGAGCGAGGAG 1820
Db |||||
1621 TGTGTTCTCAGCCAGAGAACCGGTGCGGCTGCTGGGCGCCCTGAGCTGCTGAGCGAGGAG 1680
QY 1821 GGGCGGCTTGGCTTCAAGCTTGGGAGAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1880
Db |||||
1681 GGGCGGCTTGGCTTCAAGCTTGGGAGAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740

QY 1881 AGGGAGCCAGGCGCGCGGGCTGGCTGAAGAGGCGGACTACATTGTGTCTCAGTGAATGG 1940
Db |||||
1741 AGGGAGCCAGGCGCGCGGGCTGGCTGAAGAGGCGGACTACATTGTGTCTCAGTGAATGG 1800
QY 1941 GCAGCCATGCAAGTGGTGGAGACACGCGAGGTGGTACGGAGCTGAAGGCTGCGGGAGA 2000
Db |||||
1801 GCAGCCATGCAAGTGGTGGAGACACGCGAGGTGGTACGGAGCTGAAGGCTGCGGGAGA 1860
QY 2001 GCGCGGCGCCAGCTGCGAGGTGGTCTGCTGCGCCAGCTCTAGACTTCCAGCTTGGG 2060
Db |||||
1861 GCGCGGCGCCAGCTGCGAGGTGGTCTGCTGCGCCAGCTCTAGACTTCCAGCTTGGG 1920
QY 2061 GAGCCGCGCGCCGCTGCTGCGGCGCCAGGCGGCTTCTAAGGAGCCAGAGGAGCATGG 2120
Db |||||
1921 GAGCCGCGCGCCGCTGCTGCGGCGCCAGGCGGCTTCTAAGGAGCCAGAGGAGCATGG 1980
QY 2121 TTGCAAGACCCCGGAGCTCAAGTGGGCGAGTCCCGGCGCCCTCTCAACTGAGAGCGAAA 2180
Db |||||
1981 TTGCAAGACCCCGGAGCTCAAGTGGGCGAGTCCCGGCGCCCTCTCAACTGAGAGCGAAA 2040
QY 2181 GCGCCAGCAGGCGCAAGACTGGAGGCTGCGCCAGCGCTGCGCCAGGTCAGGTCAGGCTCC 2240
Db |||||
2041 GCGCCAGCAGGCGCAAGACTGGAGGCTGCGCCAGCGCTGCGCCAGGTCAGGTCAGGCTCC 2100
QY 2241 GCGCTCATCTTGAAGCACCCAGGCTGCGGCGGCTGAGGGCGGAGTCCCTGCGAGCGCCCTCAG 2300
Db |||||
2101 GCGCTCATCTTGAAGCACCCAGGCTGCGGCGGCTGAGGGCGGAGTCCCTGCGAGCG -CCTCAG 2159
QY 2301 CCCTGGCTCAGCTGGGCAAGCACCCAGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCG 2360
Db |||||
2160 CCCTGGCTCAGCTGGGCAAGCACCCAGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCG 2219
QY 2361 CAATGCGCTGCTGCGGCTCATGCTGAGGCTGCGGCGGCTGAGGGCGGAGTCCCTGCGAGCGCCCTCAG 2420
Db |||||
2220 CAATGCGCTGCTGCGGCTCATGCTGAGGCTGCGGCGGCTGAGGGCGGAGTCCCTGCGAGCG -CCTCAG 2279
QY 2421 TGGTCAGACCTGTCTGA 2437
Db |||||
2280 TGGTCAGACCTGTCTGA 2296

RESULT 4
US-10-112-944-74
; Sequence 74, Application US/10112944
; Publication No. US20040048249A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yang, Yonghong
; APPLICANT: Weng, Gezhi
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Zhiwei
; TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
; TITLE OF INVENTION: Secreted Polypeptides
; FILE REFERENCE: 805A
; CURRENT APPLICATION NUMBER: US/10/112.944
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US 09/519,705
; PRIOR FILING DATE: 2000-03-07


```
; PRIOR APPLICATION NUMBER: US 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 09/552,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 924
; SOFTWARE: pt_FL_genes Version 5.0
; SEQ ID NO 74
; LENGTH: 2738
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (103)..(2109)
US-10-112-944-74

Query Match      84.3%; Score 2082.4; DB 13; Length 2738;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 2116; Conservative 0; Mismatches 16; Indels 4; Gaps 2;

QY      321 GGGCTGTGACTCCCTGACGAGATCCAGTGCAGTGCAGGCGCAGAGGCCGAT 380
Db      162 GGGCTGTGACTCCCTGAGCAGATCCAGTGCAGGCGCAGTGCAGAGGCCGAT 221

QY      381 TCACGAGCAGATTGACAGAGAGCTGCAGATGCGGCGGCGCTGAGAACCTCTACAGAGC 440
Db      222 TCACGAGCAGATTGACAGAGAGCTGCAGATGCGGCGGCGCTGAGAACCTCTACAGAGC 281

QY      441 CACAGCAACACCGGCTGAGAGAGACGGTTCGGCTTCGGAGCTGAGCTACGTCACATCCAA 500
Db      282 CACAGCAACACCGGCTGAGAGAGACGGTTCGGCTTCGGAGCTGAGCTACGTCACATCCAA 341

QY      501 CTTGAGCTGCTGAAGAGAGCTGAGAGAGCTCAGCGGTGAGGAGTGAACCTTCGCGCGCA 560
Db      342 CTTGAGCTGCTGAAGAGAGCTGAGAGAGCTCAGCGGTGAGGAGTGAACCTTCGCGCGCA 401

QY      561 TGGAGGAGAGCTGCTACCTGCTCCCATGATCCCTTCGGCTGAAGAGAGACCAAGAGAGCT 620
Db      402 TGGAGGAGAGCTGCTACCTGCTCCCATGATCCCTTCGGCTGAAGAGAGACCAAGAGAGCT 461

QY      621 GGAAGTGTCTACACCGCTGAAGAGAGCTGATCTCAGTGCACCTTTGGAGAGGACGCGCCTC 680
Db      462 GGAAGTGTCTACACCGCTGAAGAGAGCTGATCTCAGTGCACCTTTGGAGAGGAGCGCGCTC 521

QY      681 CTACGAGGAGAAATCAGAGAGCTGAGAGGCGCTTCGGCGAGGCCATGCGGACCCCGCCG 740
Db      522 CTACGAGGAGAAATCAGAGAGCTGAGAGGCGCTTCGGCGAGGCCATGCGGACCCCGCCG 581

QY      741 GAATGAGTTCGGGCTGAGAGCTGCTCAGAGCTATTACAAACAGCTGTGCTTCCTGGATGC 800
Db      582 GAATGAGTTCGGGCTGAGAGCTGCTCAGAGCTATTACAAACAGCTGTGCTTCCTGGATGC 641

QY      801 GCGCTTCTCACCCTCCAGAGAGCTTCGGGCTCTTCTTCACTGTGTAAGACTCGCTTAC 860
Db      642 GCGCTTCTCACCCTCCAGAGAGCTTCGGGCTCTTCTTCACTGTGTAAGACTCGCTTAC 701

QY      861 TGGGGTCCCGGCGCAGAGAGTGGCGCTTCGAGTTCGAGAGGCGAGCTTCTTCAACAT 920
Db      702 TGGGGTCCCGGCGCAGAGAGTGGCGCTTCGAGTTCGAGAGGCGAGCTTCTTCAACAT 761

QY      921 CGGTGCTTCACACGAGATTGGGCGCGCAGAGAGCTTCCTGACACGAGGCTGCGCG 980
Db      762 CGGTGCTTCACACGAGATTGGGCGCGCAGAGAGCTTCCTGACACGAGGCTGCGCG 821

QY      981 CCGGCTTATGAGAGCTTCCAGAGGCGCTTCGGGCTTCAGCTTCCTGAGGAGAACTT 1040
Db      822 CCGGCTTATGAGAGCTTCCAGAGGCGCTTCGGGCTTCAGCTTCCTGAGGAGAACTT 881

QY      1041 CTCCTATGCGGAGGCCAGACATGAGGCTTCGCTTCCTTCGCACTGGAGCAGCTCAT 1100
Db      882 CTCCTATGCGGAGGCCAGACATGAGGCTTCGCTTCCTTCGCACTGGAGCAGCTCAT 941
```

```
QY      1101 GATGGCCAGGCGCCAGGAATGTGTTTGTAGGCGCTCTTCCACCACTGCTTCCATGGCCCC 1160
Db      942 GATGGCCAGGCGCCAGGAATGTGTTTGTAGGCGCTCTTCCACCACTGCTTCCATGGCCCC 1001

QY      1161 CCAAGACTGCTGCTGGCCAGCTGCGCTGCGGAGGAGGCGCCAGGTCGAGCCAGTA 1220
Db      1002 CCAAGACTGCTGCTGGCCAGCTGCGCTGCGGAGGAGGCGCCAGGTCGAGCCAGTA 1061

QY      1221 CAGGCTAGTGTGACCGGAGACCATGSCCAGCACCCGTCACAGCTAGTGTGCTTCTCCTG 1280
Db      1062 CAGGCTAGTGTGACCGGAGACCATGSCCAGCACCCGTCACAGCTAGTGTGCTTCTCCTG 1121

QY      1281 GACTGCCCTGTGTCATGTCAAGGCGCAGTACTTCCGCTCCCTGGCCCACTACCACTAGC 1340
Db      1122 GACTGCCCTGTGTCATGTCAAGGCGCAGTACTTCCGCTCCCTGGCCCACTACCACTAGC 1181

QY      1341 CATGGCCCTGTGGAAGGCTCCCGAGGAGCGGAGGAGAGCTCCCAAGCAGCAGAGCT 1400
Db      1182 CATGGCCCTGTGGAAGGCTCCCGAGGAGCGGAGAGCTCCCAAGCAGCAGAGCT 1238

QY      1401 CTTCTGAGGCGCCCGCCACCTCTTAAGCCCGGAGGCGCTGTGCTGCTGCGCAGGAGCTGGA 1460
Db      1239 CTTCTGAGGCGCCCGCCACCTCTTAAGCCCGGAGGCGCTGTGCTGCTGCGCAGGAGCTGGA 1298

QY      1461 GGAGCGCAGGAGCTTTGGCAAGGACACCTGAAGCGTGCATCCTGGGCGAGGAGGAGGC 1520
Db      1299 GGAGCGCAGGAGCTTTGGCAAGGACACCTGAAGCGTGCATCCTGGGCGAGGAGGAGGC 1358

QY      1521 GCTGGGCTGACAGCGCTGTGCGCGCTGCTGCGGAGGAGTGGAGCTGCTTCGGGCTGTGAT 1580
Db      1359 GCTGGGCTGACAGCGCTGTGCGCGCTGCTGCGGAGGAGTGGAGCTGCTTCGGGCTGTGAT 1418

QY      1581 CTCCTGAGAGCTGACAGCGCTGCTGCTCAAGTATGCGGAGCTCGACCGTGAGGATGACTT 1640
Db      1419 CTCCTGAGAGCTGACAGCGCTGCTGCTCAAGTATGCGGAGCTCGACCGTGAGGATGACTT 1478

QY      1641 CTGTGAGGCTGCGAGGCGCCCGACATTCAGCTTAAGACCCACAGAGCCAGAGGCGCAG 1700
Db      1479 CTGTGAGGCTGCGAGGCGCCCGACATTCAGCTTAAGACCCACAGAGCCAGAGGCGCAG 1538

QY      1701 GATGCCACGCTGTCCAGGAGGAGGCGCTGACATCTTCATCGGCTGGGCGCCCTGTC 1760
Db      1539 GATGCCACGCTGTCCAGGAGGAGGCGCTGACATCTTCATCGGCTGGGCGCCCTGTC 1598

QY      1761 TGTGTTCTCAGCAAGAAACCGGTGGGCGCTGCTGAGCATCTTCATCGGCTGGGCGCCCTGTC 1820
Db      1599 TGTGTTCTCAGCAAGAAACCGGTGGGCGCTGCTGAGCATCTTCATCGGCTGGGCGCCCTGTC 1658

QY      1821 GGGCGGCTTTGGCTTACCGCTTCGGGAGAGCTGCGCTGCTCATCGCTGCGCTCATTC 1880
Db      1659 GGGCGGCTTTGGCTTACCGCTTCGGGAGAGCTGCGCTGCTCATCGCTGCGCTCATTC 1718

QY      1881 AGGAGCGAGCGCGCGCGCTGCGCTGAGGAGGCGGAGCTACATTCGTGTCAGTGAATGG 1940
Db      1719 AGGAGCGAGCGCGCGCGCTGCGCTGAGGAGGCGGAGCTACATTCGTGTCAGTGAATGG 1778

QY      1941 GCAGCAATGAGAGTGTGTGAGAGACACGCGGAGGCTGTCAGGAGCTGAAGGCTGCGGAG 2000
Db      1779 GCAGCAATGAGAGTGTGTGAGAGACACGCGGAGGCTGTCAGGAGCTGAAGGCTGCGGAG 1838

QY      2001 GGGGCGCGCAGCGCTGAGGTGTGCTGCTGCGCCAGCTCTAGACTGCGCAGCTTGGG 2060
Db      1839 GGGGCGCGCAGCGCTGAGGTGTGCTGCTGCGCCAGCTCTAGACTGCGCAGCTTGGG 1898

QY      2061 GGAACCGCGCGCGCTGCTGCTGCGCCCGCAGGCGGCTTCTAAGAGGCGCAGAGGAGCATGG 2120
Db      1899 GGAACCGCGCGCGCTGCTGCTGCGCCCGCAGGCGGCTTCTAAGAGGCGCAGAGGAGCATGG 1958

QY      2121 TTGCAAGACCCCGGCTATCCAGTGGGCGAGTCCCGCGCCCTCTCACTGAGCGCGGAAA 2180
Db      1959 TTGCAAGACCCCGGCTATCCAGTGGGCGAGTCCCGCGCCCTCTCACTGAGCGCGGAAA 2018

QY      2181 GGCCCGAGGCGCAAGACTGGAGGCTGCCCGCAGGCTGTGTGCCCGCAGTGTGAGCGAGCTCC 2240
```


Db 27 GGCAGGCAACCTGAGCGTGCATCTCTGGGAGGAGGAGCGCTTGGCTGCAGCC 86
QY 1537 CTGTGCGCGTCTCTGCGAGAGTGGACCTGCTTCCGCTGTGATCTCCAGACGCTGCAG 1596
Db 87 CTGTGCGCGTCTCTGCGAGAGTGGACCTGCTTCCGCTGTGATCTCCAGACGCTGCAG 146
QY 1597 CGCTCACTGGCCTCAAGTATGGGAGCTGCACCTGAGGATGACTTCTGTGAGGCTGCCGAG 1656
Db 147 CGCTCACTGGCCTCAAGTATGGGAGCTGCACCTGAGGATGACTTCTGTGAGGCTGCCGAG 206
QY 1657 GCCCGGACATCAGCCTTAAGACCCACAGAGCCAGAGCCAGGATGCCACGCTGTCC 1716
Db 207 GCCCGGACATCAGCCTTAAGACCCACAGAGCCAGAGCCAGGATGCCACGCTGTCC 266
QY 1717 CAGGGAGAGGGGCTGACATCTTCCATTCGCTGCGGCCCTGTCTGTGTCTCAGCCAG 1776
Db 267 CAGGGAGAGGGGCTGACATCTTCCATTCGCTGCGGCCCTGTCTGTGTCTCAGCCAG 326
QY 1777 AACGGTGGCGCTGCTGGGCGCGCTCCACTGACCCGAGGAGGCGGCTTGGCTC 1836
Db 327 AACGGTGGCGCTGCTGGGCGCGCTCCACTGACCCGAGGAGGCGGCTTGGCTC 386
QY 1837 ACGCTTCGCGGACATCGCTGTCTCATCGCTGCCCTCATTTCCAGGAGCCAGCGCGG 1896
Db 387 ACGCTTCGCGGACATCGCTGTCTCATCGCTGCCCTCATTTCCAGGAGCCAGCGCGG 446
QY 1897 GCGGCTGGCTGAAGGAGGCGCATCATTTGTGTCACTGATGAGGAGGAGGAGTGG 1956
Db 447 GCGGCTGGCTGAAGGAGGCGCATCATTTGTGTCACTGATGAGGAGGAGGAGTGG 506
QY 1957 TGGAGACACGCGAGGCTGAGGAGCTGAAGCTCGGAGAGGCGGCGCCAGCTG 2016
Db 507 TGGAGACACGCGAGGCTGAGGAGCTGAAGCTCGGAGAGGCGGCGCCAGCTG 566
QY 2017 CAGGTGTGTGCTGCTGCCAGCTCTAGACTTGCCAGCTTGGGAGCCCGCGCCGCTC 2076
Db 567 CAGGTGTGTGCTGCTGCCAGCTCTAGACTTGCCAGCTTGGGAGCCCGCGCCGCTC 626
QY 2077 CTGCTGGGCGCCAGGGGCTTCTAAGAGCCAGAGGAGGAGTGGTTGCAAGACCCCGCA 2136
Db 627 CTGCTGGGCGCCAGGGGCTTCTAAGAGCCAGAGGAGGAGTGGTTGCAAGACCCCGCA 686
QY 2137 TCACGCTGGGCGCTGCTCCCGGCGCTCTCAACTGGAGCCGAAAGCCAGCAGGCAAG 2196
Db 687 TCACGCTGGGCGCTGCTCCCGGCGCTCTCAACTGGAGCCGAAAGCCAGCAGGCAAG 746
QY 2197 ACTGGAGGCTGCCCCCAGCCCTGTGCCCCAGTGAAGCCAGCTCCGCGCTCATCTTGAAG 2256
Db 747 ACTGGAGGCTG-CCCCAGCCCTGTGCCCCAGTGAAGCCAGCTCCGCGCTCATCTTGAAG 805
QY 2257 CACCAGGCTGGCGTGAAGGCGCAGGATCCCTGACGCGCTCAGCCCTGGCTCCAGCTGG 2316
Db 806 CACCAGGCTGGCGTGAAGGCGCAGGATCCCTGACGCGCTCAGCCCTGGCTCCAGCTGG 864
QY 2317 CAGCAAGCAGCAGCAGTCCCTCCCAAGCAGGAGCTCCGCGCAATGCTCTCCGCGC 2376
Db 865 CAGCAAGCAGCAGCAGTCCCTCCCAAGCAGGAGCTCCGCGCAATGCTCTCCGCGC 924
QY 2377 TCATGCTGGAGGCTGCTTGGGCACTTGCCTGCCCCATTAAGACTGCTGAGACTGTCTG 2436
Db 925 TCATGCTGGAGGCTGCTTGGGCACTTGCCTGCCCCATTAAGACTGCTGAGACTGTCTG 984
QY 2437 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2465
Db 985 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1013

RESULT 7
US-10-023-896-43
; Sequence 43, Application US/10023896
; Publication No. US2003002776A1
; GENERAL INFORMATION:

; APPLICANT: Victor Roschke
; TITLE OF INVENTION: 29 Human Cancer Associated Proteins
; FILE REFERENCE: PA004P1
; CURRENT APPLICATION NUMBER: US/10/023,896
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: PCT/US00/23794
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/152,296
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/158,003
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 1013
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-896-43

Query Match 38.7%; Score 955.4; DB 15; Length 1013;
Best Local Similarity 99.2%; Pred. No. 5.7e-222;
Matches 981; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

QY 1477 GGCAGGCAACCTGAGCGTGCATCTCTGGGAGGAGGAGCGCTTGGCTGCAGCC 1536
Db 27 GGCAGGCAACCTGAGCGTGCATCTCTGGGAGGAGGAGCGCTTGGCTGCAGCC 86
QY 1537 CTGTGCGCGTCTCTGCGAGAGTGGACCTGCTTCCGCTGTGATCTCCAGACGCTGCAG 1596
Db 87 CTGTGCGCGTCTCTGCGAGAGTGGACCTGCTTCCGCTGTGATCTCCAGACGCTGCAG 146
QY 1597 CGCTCACTGGCCTCAAGTATGGGAGCTGCACCTGAGGATGACTTCTGTGAGGCTGCCGAG 1656
Db 147 CGCTCACTGGCCTCAAGTATGGGAGCTGCACCTGAGGATGACTTCTGTGAGGCTGCCGAG 206
QY 1657 GCCCGGACATCAGCCTTAAGACCCACAGAGCCAGAGCCAGGATGCCACGCTGTCC 1716
Db 207 GCCCGGACATCAGCCTTAAGACCCACAGAGCCAGAGCCAGGATGCCACGCTGTCC 266
QY 1717 CAGGGAGAGGGGCTGACATCTTCCATTCGCTGCGGCCCTGTCTGTGTCTCAGCCAG 1776
Db 267 CAGGGAGAGGGGCTGACATCTTCCATTCGCTGCGGCCCTGTCTGTGTCTCAGCCAG 326
QY 1777 AACGGTGGCGCTGCTGGGCGCGCTCCACTGACCCGAGGAGGCGGCTTGGCTC 1836
Db 327 AACGGTGGCGCTGCTGGGCGCGCTCCACTGACCCGAGGAGGCGGCTTGGCTC 386
QY 1837 ACGCTTCGCGGAGACTCGCCCTGTCTCATCGCTGCCCTCATTTCCAGGAGCCAGCGCGG 1896
Db 387 ACGCTTCGCGGAGACTCGCCCTGTCTCATCGCTGCCCTCATTTCCAGGAGCCAGCGCGG 446
QY 1897 GCGGCTGGCTGAAGGAGGCGCATCATTTGTGTCACTGATGAGGAGGAGGAGTGG 1956
Db 447 GCGGCTGGCTGAAGGAGGCGCATCATTTGTGTCACTGATGAGGAGGAGGAGTGG 506
QY 1957 TGGAGACACGCGAGGCTGAGGAGCTGAAGCTCGGAGAGGCGGCGCCAGCTG 2016
Db 507 TGGAGACACGCGAGGCTGAGGAGCTGAAGCTCGGAGAGGCGGCGCCAGCTG 566
QY 2017 CAGGTGTGTGCTGCTGCCAGCTCTAGACTTGCCAGCTTGGGAGCCCGCGCCGCTC 2076
Db 567 CAGGTGTGTGCTGCTGCCAGCTCTAGACTTGCCAGCTTGGGAGCCCGCGCCGCTC 626
QY 2077 CTGCTGGGCGCCAGGGGCTTCTAAGAGCCAGAGGAGGAGTGGTTGCAAGACCCCGCA 2136
Db 627 CTGCTGGGCGCCAGGGGCTTCTAAGAGCCAGAGGAGGAGTGGTTGCAAGACCCCGCA 686
QY 2137 TCACGCTGGGCGCTGCTCCCGGCGCTCTCAACTGGAGCCGAAAGCCAGCAGGCAAG 2196
Db 687 TCACGCTGGGCGCTGCTCCCGGCGCTCTCAACTGGAGCCGAAAGCCAGCAGGCAAG 746

```
QY 2197 ACTGAGGCTGCCCCAGCCCTGTGCTCCAGTGAAGCCAGCTCCGCCCTCATCTTTGAAG 2256
Db |||||
QY 747 ACTGGAGGCTG-CCCCAGGCTGTGCTCCAGTGAAGCCAGCTCCGCCCTCATCTTTGAAG 805
Db |||||
QY 2257 CACCCAGGCTGCGCTGAGGGCCAGGATCCCTGACGCCCTCAGCCCTCGCTCCAGCTGG 2316
Db |||||
QY 806 CACCCAGGCTGCGCTGAGGGCCAGGATCCCTGACG-CTCAGCCCTCGCTCCAGCTGG 864
Db |||||
QY 2317 CAGCAAGCACCCAGGATGCTCCCTCCACCCAGAGGACCTTCGGGGCAATGCTGTCCCGCC 2376
Db |||||
QY 865 CAGCAAGCACCCAGGATGCTCCCTCCACCCAGAGGACCTTCGGGGCAATGCTGTCCCGCC 924
Db |||||
QY 2377 TCATGCTGAGGCTGCTCGGGCACCTGCTGCCATTAAGACTGCTCAGACCTGTCTG 2436
Db |||||
QY 925 TCATGCTGAGGCTGCTCGGGCACCTGCTGCCATTAAGACTGCTCAGACCTGTCTG 2436
Db |||||
QY 2437 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2465
Db |||||
QY 985 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1013
Db |||||

RESULT 8
US-09-833-381-1293
; Sequence 1293, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1293
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1155)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1293
```

```
Query Match 38.3%; Score 946.8; DB 9; Length 1155;
Best Local Similarity 99.2%; Pred. No. 7.2e-220;
Matches 962; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 1477 GGCAAGGCACACCTGAAGCGTGCCATCTCGGGGAGGAGGAGCGCTCGGCTGCACGCC 1536
Db |||||
QY 133 GGCAGGACACACCTGAAGCGTGCCATCTCGGGGAGGAGGAGCGCTCGGCTGCACGCC 192
Db |||||
QY 1537 CTGTGCGCGCTCTGCGGAGGTGACCTGCTTGGGCTGTGATCTCCACAGCGTGCAG 1596
Db |||||
QY 193 CTGTGCGCGCTCTGCGGAGGTGACCTGCTTGGGCTGTGATCTCCACAGCGTGCAG 252
Db |||||
QY 1597 CGCTACTGGCCAGTATGCGGAGCTCGACCTGCGTGGAGTACTTCTGTGAGGCTGCCGAG 1656
Db |||||
QY 253 CCCCCACCGGCAAGTATGCGGAGCTCGACCTGACCGTGGAGTACTTCTGTGAGGCTGCCGAG 312
Db |||||
QY 1657 GCCCGGACATCCAGCTTAAGACCCACAGAGCCAGAGGCGAGATGCGAGCTGTGCC 1716
Db |||||
QY 313 GCCCGGACATCCAGCTTAAGACCCACAGAGCCAGAGGCGAGATGCGAGCTGTGCC 372
Db |||||
QY 1717 CAGGGAGGGGCTGACATCTTCCATCGGCTGGGGCCCTGTCTGTGTTCTCAGCCAG 1776
Db |||||
QY 373 CAGGGAGGGGCTGACATCTTCCATCGGCTGGGGCCCTGTCTGTGTTCTCAGCCAG 432
Db |||||
QY 1777 AACCGGTGGCGCTGTGGGGCCCTGACCTGACCCGAGGAGGGCGGCTTTGGCTC 1836
Db |||||
QY 433 AACCGGTGGCGCTGTGGGGCCCTGACCTGACCCGAGGAGGGCGGCTTTGGCTC 492
Db |||||
```

```
RESULT 9
US-10-112-944-558/c
; Sequence 558, Application US/10112944
; Publication No. US20040048249A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Gezhi
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Zhiwei
; TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
; FILE REFERENCE: 805A
; CURRENT APPLICATION NUMBER: US/10/112,944
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03

1837 AGCTTCGGGAGACTCGCTGTCTCTCATGCTGCGTCAATTCAGGAGCCAGGCGCG 1896
Db |||||
493 AGCTTCGGGAGACTCGCTGTCTCTCATGCTGCGTCAATTCAGGAGCCAGGCGCG 552
Db |||||
1897 GGGCTCGGCTCAAGGAGGGGCACTACATTTGTGTAGTGAATGGGAGCCATGCAAGTGG 1956
Db |||||
553 GGGCTCGGCTCAAGGAGGGGCACTACATTTGTGTAGTGAATGGGAGCCATGCAAGTGG 612
Db |||||
1957 TGGAGACACCGCGAGGTGGTGAAGGCTGGAAGGCTGCGGGAGAGCGGGCGCCAGCCTG 2016
Db |||||
613 TGGAGACACCGCGAGGTGGTGAAGGCTGGAAGGCTGCGGGAGAGCGGGCGCCAGCCTG 672
Db |||||
2017 CAGGTGGTGTCTGCTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 2076
Db |||||
673 CAGGTGGTGTCTGCTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 732
Db |||||
2077 CTGCTGGGCGCCAGGGGCTTCTTAAGAGCCAGAGGAGCATGGTTGCAAGACCCCGGCA 2136
Db |||||
733 CTGCTGGGCGCCAGGGGCTTCTTAAGAGCCAGAGGAGCATGGTTGCAAGACCCCGGCA 792
Db |||||
2137 TCCAGTGGGCGCAGTCCCGGCGCTCTCTCAACTGAGCCGAAAGCCAGCAGGCGCAAG 2196
Db |||||
793 TCCAGTGGGCGCAGTCCCGGCGCTCTCTCAACTGAGCCGAAAGCCAGCAGGCGCAAG 852
Db |||||
2197 ACTGAGGCTGCCCCAGCCCTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 2256
Db |||||
853 ACTGAGGCTGCCCCAGCCCTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 912
Db |||||
2257 CACCCAGAGGTGGCGCTGAGGGCCAGGATCCCTGACAGCCCTCAGCCCTCATCTTTGAAG 2316
Db |||||
913 CACCCAGAGGTGGCGCTGAGGGCCAGGATCCCTTGCAGC-CTCAGCCCTGGCTTCCAGCTGG 971
Db |||||
2317 CAGCAAGCACCGAGCATGCCCTCCCGCAGGAGCTTCCCGCAGTGAAGCCAGCTTCCCGCTCATCTTTGAAG 2376
Db |||||
972 CAGCAAGCACCGAGCATGCCCTCCCGCAGGAGCTTCCCGCAGGAGCTTCCCGCTCATCTTTGAAG 2436
Db |||||
2377 TCATGCTGAGGCTGCTCGGGCAGCTTCCCTGCTTCCCATTAAGACTGGTTCAGACCTGTCTG 2436
Db |||||
1032 TCATGCTGAGGCTGCTCGGGCAGCTTCCCGCAGGAGCTTCCCGCAGGAGCTTCCCGCTCATCTTTGAAG 1091
Db |||||
2437 AAAAAAAAAA 2446
Db |||||
1092 AAAAAAAAAA 1101
Db |||||
```

; PRIOR APPLICATION NUMBER: US 09/515,126									
; PRIOR FILING DATE: 2000-02-28									
; PRIOR APPLICATION NUMBER: US 09/519,705									
; PRIOR FILING DATE: 2000-03-07									
; PRIOR APPLICATION NUMBER: US 09/540,217									
; PRIOR FILING DATE: 2000-03-31									
; PRIOR APPLICATION NUMBER: US 09/552,929									
; PRIOR FILING DATE: 2000-04-18									
; PRIOR APPLICATION NUMBER: US 09/577,408									
; PRIOR FILING DATE: 2000-05-18									
; NUMBER OF SEQ ID NOS: 924									
; SOFTWARE: pt_FL_genes Version 5.0									
; SEQ ID NO 558									
; LENGTH: 1046									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
; FEATURE:									
; NAME/KEY: misc feature									
; LOCATION: (1)-(1046)									
; OTHER INFORMATION: n = a,t,c or g									
US-10-112-944-558									
Query Match									
Best Local Similarity 97.9%; Score 674; DB 13; Length 1046;									
Matches 683; Conservative 0; Mismatches 15; Indels 0; Gaps 0;									
QY	1740	CCATGGCTGGGCCCCCTGTCTGTCTCAGCCAGAACCGGTGGGGGCTGTGGGGCC	1799						
DB	702	CCCTGCCAGGGGCCCTGTCTGTCTCAGCCAGAACCGGTGGGGGCTGTGGGGCC	643						
QY	1800	CGTCCACCTCACCGAGAGGGCGCTTGGCTCAGCTTCGGGGAGACTCGCTGT	1859						
DB	642	CGTCCACCTCACCGAGAGGGCGCTTGGCTCAGCTTCGGGGAGACTCGCTGT	583						
QY	1860	CCTCATCGCTGCCCTCAATTCAGGGAGCCAGGCGCGCGCTGGCTGAAGAGGCGGA	1919						
DB	582	CCTCATCGCTGCCCTCAATTCAGGGAGCCAGGCGCGCGCTGGCTGAAGAGGCGGA	523						
QY	1920	CTACATTGTGTAGTAATGGGAGCCATGAGTGTGTGAGACACCGGAGTGTGTAC	1979						
DB	522	CTACATTGTGTAGTAATGGGAGCCATGAGTGTGTGAGACACCGGAGTGTGTAC	463						
QY	1980	GGAGCTCAAGCTCGGGAGAGGGCGGCGCAGCTCAGTGTGTGTGCTGTGCCAG	2039						
DB	462	GGAGCTCAAGCTCGGGAGAGGGCGGCGCAGCTCAGTGTGTGTGCTGTGCCAG	403						
QY	2040	CTCTAGACTGCCCCAGCTTGGGGGACCGCCGCGCCGTCTCTGTGGGCCCCAGGGGGCTTCT	2099						
DB	402	CTCTAGACTGCCCCAGCTTGGGGGACCGCCGCGCCGTCTCTGTGGGCCCCAGGGGGCTTCT	343						
QY	2100	AAGAGCCAGAGGAGCATGTTTCAAGACCCCGGCATCCAGTGGGGCAGTCCCGGCC	2159						
DB	342	AAGAGCCAGAGGAGCATGTTTCAAGACCCCGGCATCCAGTGGGGCAGTCCCGGCC	283						
QY	2160	CCTCTCAACTGGAGCCGAAAGGCCAGCAGGCAAGACTGGAGGCTGCCCCAGCCCTG	2219						
DB	282	CCTCTCAACTGGAGCCGAAAGGCCAGCAGGCAAGACTGGAGGCTGCCCCAGCCCTG	223						
QY	2220	TGCCCCAGTGAAGCAGCTCCGCTCATCTTGAAGACCCAGGTTGGCGTGAAGGCC	2279						
DB	222	TGCCCCAGTGAAGCAGCTCCGCTCATCTTGAAGACCCAGGTTGGCGTGAAGGCC	163						
QY	2280	AGGATCCCTGCAGCCCTCAGCCTGCTCCAGTGGCAGCAAGCAGCATGCCCTC	2339						
DB	162	AGGATCCCTGCAGCCCTCAGCCTGCTCCAGTGGCAGCAAGCAGCATGCCCTC	103						
QY	2340	CCCACCCAGAGACCTCCGGGCAATGCTGTCCCGCTCATCTGAGGCTGCTCGGGC	2399						
DB	102	CCCACCCAGAGACCTCCGGGCAATGCTGTCCCGCTCATCTGAGGCTGCTCGGGC	43						
QY	2400	ACCTGCCTGCCATTAAAGACTGGTTCAGACTGTCTGA	2437						
DB	42	ACCTGCCTGCCATTAAAGACTGGTTCAGACTGTCTGA	5						

QY 981 CCGCGCTATGAGGCGCTTCCAGAGGCGCGCTGGGCGCTTCAGCCTCTCGAGGAGAACTT 1040
DB 729 GAGTGCCATAGATGCTTTTCAGAGAGCGCGAGGGGTTTAAATACCTGAAACACACATT 788
QY 1041 CTCCCATGGCCGAGCCACAGACATGAGCGCTGGTCCCTCTGGCGCACTGGAGCAGCTCAT 1100
DB 789 TACCCATATCTCCAAATTACACATGAGCCCTGCCATGCTCAGCGTCTCGTCAAAATGAT 848
QY 1101 GATGGCCAGGCCACAGAAATGTGTGTTTGAAGGCGCTCTCAACACCTGCCTCCATGGCCCC 1160
DB 849 GCTTGACACAGGCCCAAGAAACGCTGTTTGAGAAAAATCAGCGCTTCCTG-----GGATCCG 902
QY 1161 CCAAGACTGCTGGCCCGCAGCTGCGCTGGCGCAGGAGCGCGCCAGGTGCGAGCCAGTA 1220
DB 903 GAATGAATCTTCATGCTGTGTAAGGTGGCTCAGAGGCGTCTTAAGGTGGAGGCTCTA 962
QY 1221 CAGGCTAGTGCACCGGACCACTGGCCCAAGCCAGCCCGTCCAGCACTTACGTGCTCTCTCTG 1280
DB 963 CCAACAGCTACACGCGCATGAGCCAGCGCGCGGTGAAGAGAAACATCCCTACTCTCTG 1022
QY 1281 GACTGCCCTGGTGTATGAGCCGAGTACTTCCGCTCCCTGGCCCACTTACCACGTAGC 1340
DB 1023 GGGCAGCTTATGCTGCTGAGGCGCCACCACTACGCGGCGCTGGCCCACTACTTCACTGC 1082
QY 1341 CATGGCCCTCTGGACGCGCTCCCGAGCGCCGAGGAGAGCTCCCGACGACAGCAGGT 1400
DB 1083 CATCTCTCTATCGACCAACAGGTGAAGCCAGGACCGATCTGGACCAACAGGAGAGTG 1142
QY 1401 CTTCTCTGCAGC---CCCCACCTCTCTTAAGCCCGAGGCGCTGTGC---TGCCGCGAGGA 1454
DB 1143 CTTCTCCAGCTCTACGACCACTGCGAGAGGCGCTGACACCTTGGCCACACTGAAGAA 1202
QY 1455 GCTGGAGAGCGAGGAGCTTGGCAAGGCACACTGAGCGTCCGATCTGGGCGAGGA 1514
DB 1203 TGATCAGCAGCGCGCAGCTGCGGGAAGTCCCACTTGGCAGAGCCATGGCTCATACGA 1262
QY 1515 GGAGCGCTGGCGTGCACGCGCTGTGCGCGGTCTCTGGCGAGGTGACCTGCTTCGGGC 1574
DB 1263 GGAGTGGTGGGAGGAGCGCTCTGCAAGAACTGGGAGCATTTGAGGTGCTACAGAA 1322
QY 1575 TGTGATCTCCAGACGCTGACGCGCTCACTGGCCAAAGTATGCGGAGCTGACCCGTGAGGA 1634
DB 1323 GGTGCTGTGTCGCGCACAGGAACGCTCCCGCTCACGTAACCGCCAGCACAGGAGGAGA 1382
QY 1635 TGACTTCTGTGAGCGTCCGAGGCGCGGACATCCAGCCTAAGACCCACAGAGCCAGA 1694
DB 1383 TGACCTGCTGAACCTGATCGACGCGCCCGCCAGTGTGTTGCTAAAACTGAGCAAGGTTGA 1442
QY 1695 GGCCAGGATGCCAGCGCTGCCAGGGGAAGGGCGCTGACATCTTCCATCGGCTGGGGCC 1754
DB 1443 CATTTATGTCCTGCTCTCCAGCTGACAGTGAAGTCTTCCAGAGCTTCTCCAGAGCTGGGCC 1502
QY 1755 CCTGCTGTGTCTCAGCCAGAAACCGGTGGCGGCTGGTGGGCGCGCTTCCACCTGACCCG 1814
DB 1503 CTTATCTGTGTTTTCGGCTAACAGCGGTGACGCGCTCTCGAAGCATCCGCTTCACTGC 1562
QY 1815 AGGAGGCGCGCTTGGCTCAGCTTCGCGGAGACTCGCTGCTCATCGCTGCGCGGT 1874
DB 1563 AGAAGAGGGGACTTGGGGTTTCACTTGAGAGGGAACCGCCCGTTAGGTTCACTTCTCT 1622
QY 1875 CATTTCCAGGAGCAGGCGCGCGGCGCTGGCGCTGAAGGAGGCGGCACTACATTTGTGTCAGT 1934
DB 1623 GATCTCTTACTGCTCTGCTCGGTGGCAGGAGCGCGGAGGAGATATATTGTTCTCCAT 1682
QY 1935 GAATGGGAGCCCATGAGGTGGTGGAGACACCGGAGGTGGTGGAGGCTGAAGGCTGC 1994
DB 1683 TCAGCTTGTGATTGTAAGTGGCTGACCGTGAAGTGAAGTATGAAGCTGCTGAAGAGCTT 1742
QY 1995 GGGAGAGCGCGCGCCAGCTGCGAGTGTGTGCTGCTG 2034
DB 1743 TGGCGAGGACGAGATCGAGATGAAGATGAAAGTCTGTAGGCTCTCTG 1782

RESULT 12
US-10-092-900A-351
; Sequence 351, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zernhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Tchernev,, Velizar T.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Liu, Yi
; APPLICANT: Anderson, David W.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Catterton, Elina
; APPLICANT: Leite, Mario W.
; APPLICANT: Zhong, Haihong
; APPLICANT: Alsobrook, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-290C
; CURRENT APPLICATION NUMBER: US/10/092,900A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: USSN 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/283,675
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: USSN 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: USSN 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/274,191
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: USSN 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: USSN 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: USSN 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: USSN 60/287,424
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 768
; SEQ ID NO 351
; LENGTH: 2109
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(1844)
US-10-092-900A-351
Query Match 20.0%; Score 494.8; DB 13; Length 2109;
Best Local Similarity 56.7%; Pred. No. 3.8e-110;

Matches	976;	Conservative	0;	Mismatches	732;	Indels	12;	Gaps	3;
QY	321	GGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGGCAGCTGCAGAGCGCGAGGCCAGAT	380						
Db	79	GGGCTGTAAATCCCTTGCACAAACCGCGCGGAGTAATTCAGAAATCAAAGAGCTGCTTT	138						
QY	381	TCACGACGAGATTGACAGGAGCTGCAGATCGGACGGCGCTGAGAACCTCTACAGAGC	440						
Db	139	GAATCAGCAGAGATCTCTGAAGACCGTSCGATGAGGACCGGACCGGAAACCTTCTGAAAGT	198						
QY	441	CACCAGCAACAACCGGCTGAGAGAGACCGTGCCTCTGAGCTGAGCTACGTCAACTCAA	500						
Db	199	GGCCACAAACTCAAAGGTGCGGAGCAAGTCGGCTGGAGCTGAGCTTCGTCAACTCAGA	258						
QY	501	CCTGCAGCTGCTGAAGGAGGAGCTGGAGAGACTCAGCGGTGGCTGGACCCCTGGCGCGCA	560						
Db	259	CCTGCAGATGCTCAAGGAAGAGCTGGAGGGGCTGAAATCTCGGTGGCGTCTATCAGAA	318						
QY	561	TGGGAGCGAAGTGTCATCTGTCCCATGATCCCCCTGGCGCTGAAGGAGCAAGAGAGCT	621						
Db	319	CACAGAGGAGGCAATTTACGATTCCTCGTATTCCTCTGGCTGAAGGAAACGAAAGACGT	378						
QY	621	GGACTTGGTCTACACCGCTGAAGGAGCTGATCTCATGTGCACTTCGAGAGAGCGGCGCTC	680						
Db	379	CGATTTTGCAGTCGTCTCAAAGGATTTATCCTCGAAATTTACACAGTGAAGATGGCTATTT	438						
QY	681	CTACGAGGCAGAAATCAGGGAGCTGGAGGCGCTCGGCAGGCGCATCGGACCCCGACCGG	740						
Db	439	ATAATGAAGATGAATTCAGATCTTATGGATCTCAGACAAGCTTGTGGAAGCTTCAGCGC	498						
QY	741	GAATGAGTCGGGCTCGAGCTGCTCAACAGCTATTACAAACAGCTGTCTTCCTGGATGC	800						
Db	499	GGATGAGCGCGGGTGGAACTGCTGATGAACATCTTCATCCAGCTGGCTTTGTGCGAG	558						
QY	801	CGCGTCTCTCACCCCTGCCAGGACCTCGGGCTCTTCTTCCACTGGTACAGCTGCTTAC	860						
Db	559	TCGATTTCTTCCCGGCCACACGGCAGATGGGACTCCTGTTCCTCTGGTATGACTCTCTCAC	618						
QY	861	TGGGGTCCCGGCCACAGAGCGTGCCTGGCCCTCGAAGGGCAGCGCTCTCTTCAACAT	920						
Db	619	CGGGTTCCGGTCAGCCAGCAGAACCTTGTCTGTGAGAGGCCAGTGTCTGTCTTCAACAC	678						
QY	921	CGTGGCCTTCAACACGAGATTGGGGCGGCCAGGACCGCTCTCTGCACCGAGGGTGCCCG	980						
Db	679	TGGGGCCCTCTACACCCAGATTGGGACCCGGTGCATCGCAGACGCAAGCTGGCTGGA	738						
QY	981	CCGCGCTATGGAGCCCTCCAGAGGCGCGCTGGGGCCTTCAGCCTCTGAGGAGAGACTT	1040						
Db	739	GAGTGCCATAGATGCCCTTCAGAGAGCGCGAGGGGTTTTAAATCTTACCTGAAGACACAT	798						
QY	1041	CTCCATGCGCCGAGCCACAGATGAGCGCTGCTGCCCTCTGCGCACTGGAGCAGCTCAT	1100						
Db	799	TACCATCTTCCAAGTTACGACATGAGCCTGCCATGCTCAGGTGCTCGTCAAAATGAT	858						
QY	1101	GATGGCCAGGCCACGGAATGTGTGTTGAGGCGCTCTCACACCTTCCTCCATGCGCCC	1160						
Db	859	GCTTGCAAAAGCCCAAGAAACGGTTTGAGAAAAATCAGCCTTCTCTG-----GGATCCG	912						
QY	1161	CCAAGACTGCTGGCCAGCTGGCGCTGGCGCAGAGGCGCCAGGTGGCAGCCGAGTA	1220						
Db	913	GAATGAATTTCTTATGCTGTGAAGTGGCTCAGAGGGCTGTAAAGTGGGAGAGTCTA	972						
QY	1221	CAGGCTAGTGCAACGGACCATGGCCCCAGCCACCGCTCACGACTAGTGTCTGTCTCTCC	1280						
Db	973	CCAACACTACACGACCCATGAGCCAGGCGCGGTGAAGAGAAATCCCTCTACTCTCTG	1032						
QY	1281	GACTGCCCTTGGTGATGTCAAGGCCGAGTACTTCCGCTCCCTGGCCCATCTACCACGTAGC	1340						
Db	1033	GGCCAGCTTAGCTGTGAAGCCCAACCACTACGCGGCCCTGGCCCACTACTTCACTGC	1092						
QY	1341	CATGGCCCTCTGGACGGCTCCCCAGCGACCGAGGGAGAGCTCCCCACGACAGCAGGT	1400						
Db	1093	CATCTCTCTCATCGACACACAGGTGAAGCCAGGCAAGGATCTGGACACACAGAGAAGTG	1152						

QY	1401	CTTCTGTGAGC-----CGCCCACTCTCTTAAGCCCCGAGGCCCTGTGTC--TGCCGACGGA	1454
Db	1153	CCTGTCCCAAGCTCTATCAGACCATATGCCAGAGGGGCTGACACCCCTTGCCACACTGAAGAA	1212
QY	1455	GCTGGAGAGCGCAGGAGCTTGCRAAGGCACACCTGAAGGTGCCATCTTGGGGAGGA	1514
Db	1213	TGATCAGACGCGCCGACAGCTTGGGGAAGTCCCACTTCGCGAGAGCCATGGCTCATCACGA	1272
QY	1515	GGAGGCGTGTGGCTGCACGCGCTGTGCCCGGCTCTCTCGCGAGGTGACCTGTCTCGGGC	1574
Db	1273	GGAGTCSGTGCGGAGGCAAGCCCTCTCGAAGAAGCTGCGGAGCATTGAGTGTCTACAGAA	1332
QY	1575	TGTGATCTCCAGACGCTGCAGCGCTCACTGGCGCAAGTATGCGAGACTGCACCGTGAGGA	1634
Db	1333	GGTGTGTGTGCGCA CAGGAAGCCTCCCGGCTACGTCAGCCCGACCCAGAGGAGGA	1392
QY	1635	TGACTTCTGTGAGGCTGCCGAGGGCCCCGGACATCCAGCCCTAAGACCCACCAGAAAGCCAGA	1694
Db	1393	TGACCTGTGAACTGTATCGACGCCCCAGTGTGTGTGTTAAACTGTGAGCAAGAGTTGA	1452
QY	1695	GGCCAGAGATGCCAGCCTGTCTCCAGGGGAAAGGGGCGCTGACATCTTCAATCGGCTTGGGGCC	1754
Db	1453	CATTATATTGCCCCAGTCTCTCCAAGCTGCAGTCAOGGACTTCTTCAGAAAGCTGGGGCCC	1512
QY	1755	CCTGTCTGTGTTCTCAGCCCAAGAACCGTTCGGCGCTGGTGGGGCCGCTCCACTGTACCCG	1814
Db	1513	CTTATCTGTGTTTCGGCTAAACAGCGGTGAGCGCCTCTCTGAAGCATCCGCTTCACTGC	1572
QY	1815	AGGAGAGGGCGGCTTTGGCCCTCA CGCTTGGGGAGACTCGCCTGTCTCTCATCGCTCCCGT	1874
Db	1573	AGAAAGGGGCACTTGGGGTTCACTTGAGAGGGAACGCCCGCTTCAGGTTCACTTCTCT	1632
QY	1875	CATTTCAGGGAGCCAGCGCGCGGGCTGGCTGAAGAGGGGCGACTACATTGTGTCAGT	1934
Db	1633	GGATCTCTTATCTGCTCTGCTCGGTGGCAGGAGCCCGGGAAGGAGATTATATTGTCTCCAT	1692
QY	1935	GAATGGGCAGCCATGCAGGTGTGTGAGACA CCGCGAGGTGGTGA CCGAGCTGAAGGCTGC	1994
Db	1693	TCAGCTTGTGATTGAATGTGGCTGACGCTGAGTGAGGTTATGAAGCTGCTGAAGAGCTT	1752
QY	1995	GGGAGAGCGGGGCCACCGCTGCAGTGTGTGCTGCTGCTG 2034	
Db	1753	TGGCGAGACGAGATCGAGATGAAGTCTGTGAGGCTCTCTG 1792	

```

RESULT 13
US-10-108-260A-2397
; Sequence 2397, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2397
; LENGTH: 3500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-2397

```

	Query Match	20.0%;	Score 494.8;	DB 16;	Length 3500;
	Best Local Similarity	56.7%;	Pred. No. 4.1e-110;		
	Matches 976;	Conservative	0;	Mismatches 732;	Indels 12; Gaps 3;
Qy	321	GGGCTGTGACTCCTTGA	CGCAGATCCAGTGGCGCCAGCTGCAGAGCCGCAGGGCCAGAT	380	
Db	105	GGGCTTAATCCCTTGC	CAAAACGGCGGAGTAAATTGCAGATCAAGAGCTGCTTTT	164	
Qy	381	TCACGACGAGATTG	CAAGGAGCTGCAGATGGGACGGCGCTGAGACCTCTTACAGAGC	440	

Db	1239	TGATCAGCAGGGCCGACAGCTGGGGAAAGTCCGACTTCGGCAGAGCCATGCTCATCACGA	1299
Qy	1515	GGAGCGCTCGGGCTGTCACGCCCTGTGTCCCGCGCTCTGTGGCGAGGTGACCTGTCTCGGGC	1574
Db	1299	GGATTCGGTGGGGAGCGAGCCTCTGCCAAGAGCTTCGGGAGCATTTGAGGTGCTTACAGAA	1358
Qy	1575	TGTGATCTCCACAGAGCTGTGAGGCGCTCACTTGGCCAAAGTATGCGGAGGCTCGACCGTGAAGGA	1634
Db	1359	GGTGCTGTGTGGCGCACAGGAAAGCTCCCGGCTCACGTACGCCGAGCACCGAGGAGGAGGA	1418
Qy	1635	TGACTTCTGTGAGGCTCCGAGGCCCCGGACATCCAGCCCTAAGACCCACACAGAGCCAGA	1694
Db	1419	TGACTCTCTGAACCTGATCGACGCCCCCAGCTGTGTTGCTTAAACTGAGCAAGAGGTTGA	1478
Qy	1695	GGCCAGGATGCCAGCGCTGTCCACAGGGAAAGGGCCCTGAATCTTCCATCGGCTGGGGCC	1754
Db	1479	CATTATATTTGCCCCAGTTTCTCCAAAGCTGACAGTCACGACATTTCTTCCAAAGCTGGGCC	1538
Qy	1755	CCTGTCGTGTGTTCTCAGCCCAAGAACCGGTGGCGGCTGGTGGGGCCCGTCCACTGACCCG	1814
Db	1539	CTTATCTGTGTTTTTCGGCTAACAAAGCGGTGGACGCTCTCGAAGCATCGCTTCACTGC	1598
Qy	1815	AGGAGAGGGCGGCTTTGGCTCAGCTTCGGGAGACTCGCGCTGTCTCATCGCTGGCGGT	1874
Db	1599	AGAAGAAAGGAGACTTGGGGTTACCTTTGAGAGGAAGCGCCCCGTTGAGGTTCACTTCTCT	1658
Qy	1875	CATTCCAGGGAGCCAGCGCGCGGCTTGGCCTGAAGAGAGGGGACATTTGTTGTCAGT	1934
Db	1659	GGATCCTTTACTGCTCTGCTTCGGTGGCAGGAGCCCGGAGAGGAGATTATTTGTTCTCAT	1718
Qy	1935	GAATGGGCGGCATTCGAGGTGGTGGACACCGCGAGGTGGTGAACGGAGGCTGGAAGGCTGC	1994
Db	1719	TCAGCTTGTGATTGTAAATGGCTGACGCTGAGTGAGTTATGAAGCTGCTGAAGAGCTT	1778
Qy	1995	GGGAGAGCGGGCCGCCACTGCAAGTGGTGGTTCGCTGCTG	2034
Db	1779	TGGCGAGGACGAGATCGAGATGGAAGTCGTGAGGCTCTCTG	1818

RESULT 14

```

US-09-895-040A-2
; Sequence 2, Application US/09895040A
; Patent No. US20020123474A1
; GENERAL INFORMATION:
; APPLICANT: Shannon, Mark
; APPLICANT: Ji, Yonggang
; TITLE OF INVENTION: HUMAN GTP-RHO BINDING PROTEIN 2
; FILE REFERENCE: AEMICA-11
; CURRENT APPLICATION NUMBER: US/09/895,040A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: Aemica Sequence Listing Engine
; SEQ ID NO 2
; LENGTH: 2061
; TYPE: DNA

```

[illegible]

; ORGANISM: Homo sapiens
US-09-895-040A-2

Query Match 20.0%; Score 493.2; DB 9; Length 2061;
Best Local Similarity 56.7%; Pred. No. 9.2e-110;
Matches 975; Conservative 0; Mismatches 733; Indels 12; Gaps 3;

```
QY 321 GGGCTGTGACTCCCTGACGAGATCCAGTGTGGCGGCGAGCTGACAGCGCGAGGGCCAGAT 380
D 69 GGGCTGTAAATCCCTTGCACAAACCGCGCGAGTAAATTGCAGATCAAAAGAGCTGCTT 128
QY 381 TCACACGAGATGTACAGAGCTGCAGATGCGGACGGCGCTGAGAACTCTACAGAGC 440
D 129 GAATCAGCAGATCCTGAAGCCGTGCGATGAGATCGGAGCGGAAACCTTCTGAAAGT 188
QY 441 CACCGACAAACACCGGCTGAGAGAGCGGTGCGCCCTGGAGCTGAGTCAAGTCCCAA 500
D 189 GGGCCACAAACTCAAGGTCGGGAGCAAGTGGCGCTGGAGCTGAGCTTCTGCACTCAGA 248
QY 501 CCTGACGCTGCTGAAGGAGGCTGAGAGCTCAGCGGTGGGTGGAACCTGCGCGGCA 560
D 249 CTTGACAGATGCTCAAGGAAGAGCTGAGGGGCTGAAACATCTCGTGGGCTCTATCAGAA 308
QY 561 TGGGAGCGAAGCTGTCACTGTCCCAATGATCCCTCGGCGCTGAGGAGCAAGAGCT 620
D 309 CACAGAGGAGCATTTACATTTCCCTGATTTCTCTTGGCCCTGAGGAAACGAAAGAGCT 368
QY 621 GAAGTGTCTACACCGCTGAAGAGCTGATCTCAGTGCACTTTTGGAGAGGACGGCGCTC 680
D 369 CGACTTTGAGTCTGCTCAAGGATTTTATCTGGAACATTACAGTGAAGATGGCTATT 428
QY 681 CTACGAGGAGCAAAATCAGGAGCTGAGGCGCTTGGGAGGCGCATGCGGACCCCGAGCG 740
D 429 ATATGAAGATGAAATGCAAGATCTTATGATCTGAGACAAGCTGTGCGAGCGCTAGCGG 488
QY 741 GAATGAGTGGGCGCTGGAGCTGTCTACAGCTATTACACAGCTGTGCTTCTGATGTC 800
D 489 GGATGAGGCGGGGTGGAAGTCTGATGATACATTTATCCAGCTGGGCTTTTGTGAGAG 548
QY 801 GCGCTTCTCAACCCCTGCGAGAGCTCTGGGCTCTTTCTTCCACTGTGTCAGACTCGCTTAC 860
D 549 TCGATTTCTCCGCGCCACACGCGAGATGGACTCTCTGTTTCACTGGTATGATCTCTCAC 608
QY 861 TGGGCTCCGCGCCAGCAGCTGCGCTGCTTGGAGAGGCGAGCTTCTTCTCAACAT 920
D 609 CGGGTTCGGTCAACCGACAGAACCTGCTGTGAGAGAGGCGAGTGTCTCTTCTCAACAC 668
QY 921 CGGTGCTCTCACACGACAGATTTGGGCGCGCGCGCAGGACGCTCTCTCACCGAGGCTGCCG 980
D 669 TGGGCGCTCTACACCGAGATTTGGAGCCCGGTGGATCGGAGACGAGCTGGGCTGGA 728
QY 981 CCGGCTATGAGGCGCTTCCAGAGGCGCTGCGGCGCTTCAAGCTCTTGGAGGAGAACTT 1040
D 729 GAGTCCCATAGATGCTTTACAGAGAGCGCGAGGGGTTTTTAATTAACCTGAAGACACATT 788
QY 1041 CTCACATGGCGGCGCCAGACATGAGCTGCTGCTTCTGCGACTGGAGCAGCTCAT 1100
D 789 TACCCATATCTCAAGTTACGATGAGCCCTGCCATGCTCAGCGTGTCTCGTCAAAATGAT 848
QY 1101 GATGCGCCAGGCGCCAGGAATGTGTGAGGCGCTCTCACACCTGCTCCATGCGCCC 1160
D 849 GCTTGCAACAGCCCAAGAAAGGTGTTGAGAAATCAGCTTCTCTG-----GGATCCG 902
QY 1161 CCAAGACTGCTGGCGCCAGCTGCGCTGCGCAGGAGCGCGCCAGGTGGCAGCGAGTA 1220
D 903 GAAATGAAATTTCTATGCTGTTGAAGGTGGCTCAGGAGGCTGTAAAGTTGGAGAGGTCTA 962
QY 1221 CAGGCTAGTGACCGGACCATGCGCCAGCCCGCTCAGAGCTACAGTACGTGCTCTCTG 1280
D 963 CCAACAGCTACACGAGCCATGAGCGCGCGGTGAAAGAGAACATCCCTACTCTCTG 1022
QY 1281 GACTGCCCTGGTGCATGTCAGGCGGAGTACTTCCGCTCCCTGGCGCACTACAGTAGC 1340
```

```
D 1023 GGCCAGCTTTAGCTGCTGTAAGGCCCAACCACTAAGCGGCGCTGCGCCACTTACTTCACTGC 1082
QY 1341 CATGGCCCTCTGCGACGGTCCCCACGACCGAGGAGAGAGCTCCCAACGACGAGCAGGT 1400
D 1083 CATCTCTCTATGACACCAACAGGTGAAGCCAGGACGAGTCTGGACCAACGAGGAGAGTG 1142
QY 1401 CTTCTGTGACG-----CCCCACCTCTTAAGCCCGAGGCGCTGTGC--TGCCGACGGA 1454
D 1143 CTTGTCCAGCTCTACGACCAATGCAAGGGGCTGACACCTTGGCCACACTGAAGAA 1202
QY 1455 GCTGGAGGAGCGGACGAGCTTTGGCAAGGCAACCTGAAGCGTGCCTCTCTGGGGGAGGA 1514
D 1203 TGATCAGACGCGCGCAGCTGCGGGAAGTCCCACTTGCGCAGAGCCATGGCTCATCAGCA 1262
QY 1515 GGAGGCGCTGCGGCTGTCACGCGCTGTGCGCGCTCTGCGCGAGGTGGACCTGCTTCGGGC 1574
D 1263 GAGTCTGCTGCGGAGGCAAGCTCTGCAAGAGCTGCGAGCATTTGAGGTGCTACAGAA 1322
QY 1575 TGTGATCTCCAGACGCTGTCAGCGCTCTACTGCGCCAAAGTATGCGGAGCTCGACCGTGAGGA 1634
D 1323 GGTGCTGTGCGCGCACAGGAACGCTCCCGGCTCAGTACGCTGACAGCTTCTTCCAGAGCTGGGCC 1502
QY 1635 TCACTTCTGTGAGGCTGCGGAGCGCGCGGACATCCAGCTTAAGACCCACGAGCCAGA 1694
D 1383 TGACCTGTGAAACCTGATCGACGCGCGCCAGTGTGTTGCTTAAACTGAGCAAGAGGTTGA 1442
QY 1695 GGCCAGGATGCCACGCTGTCCACGAGGGAAGGGCTGACATCTTCCATCGGTGGGGCC 1754
D 1443 CATTATATGCCCCAGTCTTCCAGCTGACATGACAGCTTCTTCCAGAGCTGGGCC 1502
QY 1755 CTTGTCTGTGTTCTCAGCAAGAACCGGTGCGGCTGTGTTGGGCGCGCTCCACTGACCGG 1814
D 1503 CTTATCTGTGTTTTCGGCTTAAACAGCGGTGAGCGCTCTCTCGAAGCATCCGCTTCACTGC 1562
QY 1815 AGGAGGCGGCTTTTGGCTCAGCGCTTCCGCGGAGCTGCGCTGCTCTCATCGTGGCGT 1874
D 1563 AGAAGAGGAGGACTTTGGGGTTTCCACTTGAGAGGGAACGCGCCCGTTTCAAGTTTCACTTCT 1622
QY 1875 CATTTCAGGAGCGGCGCGCTGCGCTGAGGAGGAGGCGGAGCTACATTTGTGTCACT 1934
D 1623 GGATCTTACTGCTCTGCTGCTGCGTGGAGAGCGCGGAGAGAGATTAATTTGTCTTCAT 1682
QY 1935 GAATGGCAGCAATCAGGTGTTGGAGACACGCGGAGGTGGTGACGAGGCTGAAGGCTGC 1994
D 1683 TCAGCTTGTGATTTGTAAGTGGCTGACGCTGAGTGAAGTTATGAAGCTGCTGAAGAGCTT 1742
QY 1995 GGGAGGCGGCGCGCAGCTGACGCTGCGTGTGCTGCTGCTG 2034
D 1743 TGGCAGGAGCGAGATCGAGATGAAGTCTGTAGAGCTCTCTG 1782
```

RESULT 15

US-09-895-040A-1
; Sequence 1, Application US/09895040A
; Patent No. US20020123474A1
; GENERAL INFORMATION:
; APPLICANT: Shannon, Mark
; APPLICANT: Ji, Yonggang
; TITLE OF INVENTION: HUMAN GTP-RHO BINDING PROTEIN 2
; FILE REFERENCE: AEMICA-11
; CURRENT APPLICATION NUMBER: US/09/895,040A
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: Aemilica Sequence Listing Engine
; SEQ ID NO 1
; LENGTH: 3484
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-895-040A-1

Query Match      20.0%; Score 493.2; DB 9; Length 3484;
Best Local Similarity 56.7%; Pred. No. 9.9e-110;
Matches 975; Conservative 0; Mismatches 733; Indels 12; Gaps 3;

QY 321 GGGCTGTGACTCCCTGACGAGATCCAGTGGCGCAGCTGCAGAGCGCGAGGCCCCAGAT 380
Db |||||
QY 89 GGGCTGTGAAATCCCTTGCACAAACCGCGCGAGTAAATTCAGAAATCAAAGAGCTGCTT 148
Db |||||
QY 381 TCACACGACAGATTGACAAAGAGCTGCAGATCGGACGGGGCTGAGAACCTCTACAGAGC 440
Db |||||
QY 149 GAATCAGCAGATCTTGAAGCCGTGCGATGAGGATCGGAGCGGAAACCTTCTGAAAGT 208
Db |||||
QY 441 CACACGACAAACCGGCTGAGAGAGACGGTGGCCCTGGAGTGAAGTACGTCAACTCAA 500
Db |||||
QY 209 GGGCACAACTCAAAGCTGCGGAGCAAGTGGCGCTGAGCTGAGCTTCTGCAACTCAGA 268
Db |||||
QY 501 CTTGACCTCTGAAGGAGCTGAGAGAGCTCAGCGGTGGCTGAGCCCTGGCGGCA 560
Db |||||
QY 269 CTTGACAGATGCTCAAGGAGAGCTGGAGGGGCTGAACATCTCGGTGGCGCTGATCAGAA 328
Db |||||
QY 561 TGGGAGCGAAGCTGCTCACTGTCCTCCATGATATCCCTCTGGGCTGAAGGAGACCAAGAGCT 620
Db |||||
QY 329 CACAGAGGAGCAATTACGATTCCTCTGATTCCTCTGGCTGAAGGAAACGAAAGAGCT 388
Db |||||
QY 621 GGACTGTGCTACACGGCTGAAGGAGCTGATCTCAGTGCACTTTGGAGAGAGCGGCTC 680
Db |||||
QY 389 CGACTTTGCACTGCTCTCAAGGATTTTATCTGGAACATTTACAGTGAAGATGSGCTATT 448
Db |||||
QY 681 CTACGAGCAGAAATCAGGAGCTGAGGCGCTCGGCGAGGCCATCGGACCCCGAGCG 740
Db |||||
QY 449 ATATGAAGATGAAATTCAGATCTTATGGATCTGAGACAACTGTCGAGCGCTTAGCCG 508
Db |||||
QY 741 GAATGAGTCGGGCTGAGCTGCTCACAGCCTATTACAACCAAGCTGTGCTTCTTGGATGC 800
Db |||||
QY 509 GGATGAGGCGGGGTGGAATGCTGATGACATACTTCATCCAGCTGGGCTTTGTCGAGAG 568
Db |||||
QY 801 GCGCTTCTACCCCTGCGAGAGCTCGGGCTCTTCTTCACTGGTACGACTCGCTTAC 860
Db |||||
QY 569 TCGATTTCTTCCGCGCCACACGCGCAGATGGGACTCTCTGTTCACCTGGTATGACTCTCT 628
Db |||||
QY 861 TGGGCTCCCGCCAGCAGCTGCGCTTCGAGAGGCGAGCGTTCCTCTTCAACAT 920
Db |||||
QY 629 CGGGGTTCCCGTACGACAGAGAACTGCTGCTGAGAGAGCCAGTGTCTCTTCAACAC 688
Db |||||
QY 921 CGGTGCCCTCCACACGAGATTGGGGCGCGCAGGACCGCTCTCTGCAACCGAGGTGCGCG 980
Db |||||
QY 689 TGGGGCCCTCTACACCCAGATTGGGACCCGCTGCGATCGGCAGACCGAGGCTGGGCTGA 748
Db |||||
QY 981 CCGCGCTATGAGGCTTCCAGAGGGCGCTGGGCGCTTACGCTCTCTGAGGGAGAACTT 1040
Db |||||
QY 749 GAGTGCCATAGATGCTTTCAGAGAGCGCGAGGGGTTTAAATTTACCTGAAAGACACATT 808
Db |||||
QY 1041 CTCCTCATGCGGAGCCAGACATGAGCGCTGCTGCTCTCTGCGCACTGGAGCAGCTCAT 1100
Db |||||
QY 809 TACCCATACCTCAAGTTACGACATGAGCCCTGCCATGCTCAGCGTGTCTCTGCTCAAATGAT 868
Db |||||
QY 1101 GATGGCCCGAGCCCGAGGAATGTGTTGAGGGCGCTCTCAACACTGCTCTCAATGGGCCCC 1160
Db |||||
```

```
869 GCTTGCACAGCCCAAGAAAGCGTGTGAGAAAAATCAGCCTTCTCTG-----GGATCCG 922
1161 CCAAGACTGCTGGCCAGCTGCGCTGGCGCAGAGAGCGCCAGCTGCGAGCCGAGTA 1220
923 GAATGAATCTTCATGCTGGTGAAGTGGCTCAGGAGGCTGTGAAGTGGAGAGTCTA 982
1221 CAGGCTAGTGCACCGGACCATGGCCAGCCACCGCTCCACGACTACGTGCTCTCTCT 1280
983 CCAACAGCTACACGAGCCATGAGCCAGGCGCGGTGAAGAGAACATCCCTACTCTCT 1042
1281 GACTGCCCTGGTGCAATGTCAAGCCGAGTACTTCGGCTCCCTGGCCCACTACACAGTAGC 1340
1043 GGGCAGCTTAGCCTGCTGAAGGCCCACTACGAGGCGCTGGCCCACTACTTCACTGC 1102
1341 CATGGCCCTCTGGAGCGGCTCCCGACGACCGAGGAGAGCTCCCGACGACGAGCAGGT 1400
1103 CATCTCTCTCATCGACCACCAAGTGAGCCAGGACCGGATCTGACACACAGAGAAAGTG 1162
1401 CTTCTCTGAGC-----CCCCCACCCTCTTAAGCCCCGAGGCCCTCTGTGC--TGCGCGCAGGA 1454
1163 CCGTGTCCAGCTCTACGACCACATGCCAGAGGGGCTGACACCTTGGCCACACTGAAGAA 1222
1455 GCTGAGAGGCGCAGGAGCTTGGCAAGGACACCTGAAAGCGTGGCCATCTGGGGCAGGA 1514
1223 TGATCAGCAGCGCGCAGACTGGGGAAGTCCCACTTGGCGAGAGCCATGGCTCATACGA 1282
1515 GGAGGCGCTCGGCTGCACGCCCTGTGCCGCTCTCGCGAGGTGACCTGCTTCGGGC 1574
1283 GAGTGGCTGGGAGGCAAGCTCTGCAAGAGCTGCGGAGCATTGAGGTGCTACAGAA 1342
1575 TGTGATCTCCACAGACGCTGCAGCGCTCACTGGGCCAAGTATTCGGAGGTCCGACCGTAGGA 1634
1343 GGTGCTGTGTGCGCACAGGAACGCTCCCGCTCACGTACGCTCCCGCAGCACAGGAGGAGGA 1402
1635 TGACTTCTGTGAGCTGCCAGGCGCCCGGACATCCAGCTTAAGACCCACAGAGGCCAGA 1694
1403 TGACTTCTGAAACCTGATTCGACGCGCCCAAGTGTGTTGTTGCTAAAACTGAGCAAGAGTTGA 1452
1695 GGCCAGGATGCAAGCTGTCCAGGGGAAAGGGCCCTGACATCTTCCATCGGCTGGGGCC 1754
1463 CATTAATTGCCCCAGTTCCTCAAGCTGACAGTCAAGTCACTTCTTCCAGAGCTGGGCC 1522
1755 CCGTCTGTGTTCTCAGCCAAAGAACCGGTGGCGGCTGGTGGGGCCCGTCCACCTGACCCG 1814
1523 CTTATCTGTGTTTTTCGGCTAACAAAGCGTGGACGCTCTCTCGAAGCATCCCGCTTCACTGC 1582
1815 AGGAGAGGCGGCTTTGGCCCTCAGCTTCGGGAGACTCGGCTCTCTCTCATCGCTGCCGT 1874
1583 AGAAGAGGGGAGCTTGGGGTTTCACTTGAAGAGGAAACGCCCCCGTTTCAAGTTCACTTCT 1642
1875 CATTCAGGAGAGCCAGCGCGCGCTGGGCTGGAAGGAGGGCGACTACATTTGTGTCACT 1934
1643 GGATCTTACTGCTCTGCTCGGTGGCAGGAGCCCGGAGAGAGATATATTGTCTCCAT 1702
1935 GAATGGGACGCAATGCAAGGTGGTGAGACACGCGGAGGTGGTGGAGGAGGTGAGAGGCTGC 1994
1703 TCAGCTGTGATTGTAAGTGGCTGACGCTGAGTGAGGTTATGAAGTGTGCTGAAGAGCTT 1762
1995 GGGAGAGGCGGCGCGAGCTGACAGGTGGTGGCTGCTG 2034
1763 TGGCGAGGACGAGATCGAGATGAAAGTGGTGGCTCTG 1802
```

Search completed: July 13, 2004, 12:02:39
Job time : 1011.06 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 13, 2004, 00:06:40 ; Search time 122.439 Seconds
(without alignments)
11190.705 Million cell updates/sec

Title: US-10-697-266-1

Perfect score: 2469

Sequence: 1 tcgcggccagtggtgcgg.....aaaaaaaaaaaaaaaaa 2469

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*

2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*

3: /cgn2_6/prodata/2/ina/6A_COMB.seq:*

4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*

5: /cgn2_6/prodata/2/ina/PCITUS_COMB.seq:*

6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2469	100.0	2469	4	US-09-849-334-1	Sequence 1, Appli
2	2469	100.0	2469	4	US-10-274-878-1	Sequence 1, Appli
3	946.8	38.3	1155	4	US-09-833-381-1293	Sequence 1293, Ap
4	368	14.9	19025	4	US-09-849-334-3	Sequence 3, Appli
5	368	14.9	19025	4	US-10-274-878-3	Sequence 3, Appli
6	68	2.8	2469	4	US-09-849-334-1	Sequence 1, Appli
7	68	2.8	2469	4	US-10-274-878-1	Sequence 1, Appli
8	57.6	2.3	38653	4	US-09-922-445-1	Sequence 1, Appli
9	57.2	2.3	28958	1	US-08-258-261B-6	Sequence 6, Appli
10	57.2	2.3	28958	1	US-08-456-837-6	Sequence 6, Appli
11	57.2	2.3	28958	1	US-08-457-342-6	Sequence 6, Appli
12	57.2	2.3	28958	1	US-08-457-646A-6	Sequence 6, Appli
13	57.2	2.3	28958	1	US-08-458-076A-6	Sequence 6, Appli
14	57.2	2.3	28958	1	US-08-764-233A-4	Sequence 4, Appli
15	57.2	2.3	28958	1	US-08-457-335A-6	Sequence 6, Appli
16	57.2	2.3	28958	1	US-08-729-214-6	Sequence 6, Appli
17	57.2	2.3	28958	3	US-09-028-934-6	Sequence 6, Appli
18	57.2	2.3	49377	1	US-08-764-233A-1	Sequence 1, Appli
19	54.6	2.2	7218	1	US-08-232-463-14	Sequence 14, Appl
20	52	2.1	1578	4	US-09-252-991A-5965	Sequence 5965, Ap
21	52	2.1	2469	4	US-09-252-991A-5908	Sequence 5908, Ap
22	52	2.1	3135	4	US-09-252-991A-5922	Sequence 5922, Ap
23	51.8	2.1	7065	4	US-09-874-923-115	Sequence 115, App
24	51.2	2.1	1155	4	US-09-833-381-1293	Sequence 1293, Ap
25	51	2.1	9053	4	US-09-976-594-306	Sequence 306, App
26	49.8	2.0	1992	4	US-09-252-991A-13771	Sequence 13771, A
27	49.8	2.0	2175	4	US-09-252-991A-13657	Sequence 13657, A

ALIGNMENTS

RESULT 1

US-09-849-334-1

; Sequence 1, Application US/09849334

; Patent No. 6500655

; GENERAL INFORMATION:

; APPLICANT: RUSCH, Douglas et al

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; FILE REFERENCE: THERIOF

; FILE REFERENCE: CL001099-CIP

; CURRENT APPLICATION NUMBER: US/09/849,334

; CURRENT FILING DATE: 2001-05-07

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 2469

; TYPE: DNA

; ORGANISM: Human

US-09-849-334-1

Query Match

Best Local Similarity 100.0%; Score 2469; DB 4; Length 2469;

Matches 2469; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TCGCGCCAGTGTGTGCGGGCGGCCCTAGCCGCTGCGGAGCGCTGCGGAGCGCGG	60
DB	1	TCGCGCCAGTGTGTGCGGGCGGCCCTAGCCGCTGCGGAGCGCTGCGGAGCGCGG	60
QY	61	GCTGCTGACCCCGAGGACCCCGAGGAGCGCGGCTGCGGAGCGCTGCGGAGGAGG	120
DB	61	GCTGCTGACCCCGAGGAGCCCGAGGAGCGCGGCTGCGGAGCGCTGCGGAGGAGG	120
QY	121	CCGACGCGCGCGCGCGCGCGAGGAGCGCGGCTGCGGAGCGCTGCGGAGGAGGAGG	180
DB	121	CCGACGCGCGCGCGCGCGCGAGGAGCGCGGCTGCGGAGCGCTGCGGAGGAGGAGG	180
QY	181	AGGAAACACGTGTGAGCTCTTTACGGGGAAGAGCGGAGGCGCTGAGAGAGCTGTGCG	240
DB	181	AGGAAACACGTGTGAGCTCTTTACGGGGAAGAGCGGAGGCGCTGAGAGAGCTGTGCG	240
QY	241	TGGAGAGGCTGCGGCTTCCAGAGGGAAGACCCAGTGTGTGACAGCTTGCCCATG	300
DB	241	TGGAGAGGCTGCGGCTTCCAGAGGGAAGACCCAGTGTGTGACAGCTTGCCCATG	300
QY	301	AATCGGAGCTTATGACGTGGCTGTGACTCCCTGACGAGATCCAGTGGCGGCGCTG	360
DB	301	AATCGGAGCTTATGACGTGGCTGTGACTCCCTGACGAGATCCAGTGGCGGCGCTG	360
QY	361	CAGAGCGGCGGCGGCGGCGGAGTTCACGAGATTCAGAGGAGCTGCGAGATCGGAGG	420

361 CAGAGCCGAGGGCCAGATTACACAGCAGATTGACAGGAGCTGCAGATGGCGGCG 420
421 GCTGAGAACTCTACAGAGCCACAGGACCAACCCGGGTGAGAGAGAGGTGCGCCCTGGAG 480
421 GCTGAGAACTCTACAGAGCCACAGCACAACCCGGGTGAGAGAGAGGTGCGCCCTGGAG 480
481 CTGAGCTAGCTCAACTCCCACTGAGCTGCTCAAGGAGGAGCTGAGGAGGCTCAGCGGT 540
481 CTGAGCTAGCTCAACTCCCACTGAGCTGCTCAAGGAGGAGCTGAGGAGGCTCAGCGGT 540
541 GCGGTGAGACCTGGCGGCACTGGAGCGAAGCTGTCACTGTCCCCATGATFCCCCCTGGGC 600
541 GCGGTGAGACCTGGCGGCACTGGAGCGAAGCTGTCACTGTCCCCATGATFCCCCCTGGGC 600
601 CTGAAGAGACCAAGGAGCTGGAGCTGTCTACACCGCTGAAGGAGCTGTCTCAGTGCCAC 660
601 CTGAAGAGACCAAGGAGCTGGAGCTGTCTACACCGCTGAAGGAGCTGTCTCAGTGCCAC 660
661 TTTTGGAGAGACCGCGCTCTCTACGAGGAGAAATCAGGAGCTGGAGGCTTGGCGCAG 720
661 TTTTGGAGAGACCGCGCTCTCTACGAGGAGAAATCAGGAGCTGGAGGCTTGGCGCAG 720
721 GCCATGCGGACCCCAAGCGGAAATGAGTCGGGCTGGAGCTGTCTCAAGGCTATTACAAC 780
721 GCCATGCGGACCCCAAGCGGAAATGAGTCGGGCTGGAGCTGTCTCAAGGCTATTACAAC 780
781 CAGCTGTGCTTCTGAGTGGCGCTTCTTCAAGGCTGGAGCTGTCTCAAGGCTATTACAAC 840
781 CAGCTGTGCTTCTGAGTGGCGCTTCTTCAAGGCTGGAGCTGTCTCAAGGCTATTACAAC 840
841 CACTGTGACACTGCTTACTTGGGTCCTCCGCGCCAGCAGGCTGCCCTGGGCTTCGAGAAG 900
841 CACTGTGACACTGCTTACTTGGGTCCTCCGCGCCAGCAGGCTGCCCTGGGCTTCGAGAAG 900
901 GGCAGCGTTCTCTTCAACATCGGTGCTTCCACACGAGATTGGGGCGCGCAGGACCGC 960
901 GGCAGCGTTCTCTTCAACATCGGTGCTTCCACACGAGATTGGGGCGCGCAGGACCGC 960
961 TCCTGCAACGAGGTGCGCGCGCTTATGGAGCCTTCCAGAGGCGCGTGGGCGCTTC 1020
961 TCCTGCAACGAGGTGCGCGCGCTTATGGAGCCTTCCAGAGGCGCGTGGGCGCTTC 1020
1021 AGCCTCTGAGGAGAACTTCTCCATGCGCGGAGCCAGACATGAGCGCTGCGTCCCTC 1080
1021 AGCCTCTGAGGAGAACTTCTCCATGCGCGGAGCCAGACATGAGCGCTGCGTCCCTC 1080
1081 TGCGCACTGAGGAGCTCATGATGGCCAGGCGCCAGGAAATGTGTGTTTGGGGCTCTCA 1140
1081 TGCGCACTGAGGAGCTCATGATGGCCAGGCGCCAGGAAATGTGTGTTTGGGGCTCTCA 1140
1141 CCACCTGCTCCATGGCCCGCCAGAGCTGCTGGCCCGCAGCTGCGCTGGCGCAGGAGCC 1200
1141 CCACCTGCTCCATGGCCCGCCAGAGCTGCTGGCCCGCAGCTGCGCTGGCGCAGGAGCC 1200
1201 GCCCAGGTGGCAGCCGAGTACAGGCTAGTGACCGGACCATGGGCCAGCCACCGCTCCAC 1260
1201 GCCCAGGTGGCAGCCGAGTACAGGCTAGTGACCGGACCATGGGCCAGCCACCGCTCCAC 1260
1261 GACTAGGTGCTGCTCTGAGCTGCGCTGTGTGATGATCAAGGCGGAGTACTTCCGCTCC 1320
1261 GACTAGGTGCTGCTCTGAGCTGCGCTGTGTGATGATCAAGGCGGAGTACTTCCGCTCC 1320
1321 CTGGCCCACTACAGGAGCTGCTTCTGAGCGGCTCTCCAGCGGCTCCCGAGCGAGGAGAG 1380
1321 CTGGCCCACTACAGGAGCTGCTTCTGAGCGGCTCTCCAGCGGCTCCCGAGCGAGGAGAG 1380
1381 CTCCCCAGCAGGAGCTTCTCTGAGCGGCTCTCTGAGCGGCTCTCTTAAAGCGGCGCT 1440
1381 CTCCCCAGCAGGAGCTTCTCTGAGCGGCTCTCTGAGCGGCTCTCTTAAAGCGGCGCT 1440
1441 GTGCTGCCGAGGAGCTGGAGGAGCGGAGCTTGGCAGGAGCTTGGCAGGAGCTTGGCAGGAGCT 1500
1441 GTGCTGCCGAGGAGCTGGAGGAGCGGAGCTTGGCAGGAGCTTGGCAGGAGCTTGGCAGGAGCT 1500

1501 ATCTCTGGCAGAGAGGCGCTGCGGCTGCAGCGCCCTGTGCGCGCTCTCTGCGCGAGGTG 1560
1501 ATCTCTGGCAGAGAGGCGCTGCGGCTGCAGCGCCCTGTGCGCGCTCTCTGCGCGAGGTG 1560
1561 GACCTCTTGGGCTGTGATCTCCAGACGCTGCGAGGCTCACTGGCACAAGTATGCGGAG 1620
1561 GACCTCTTGGGCTGTGATCTCCAGACGCTGCGAGGCTCACTGGCACAAGTATGCGGAG 1620
1621 CTCGACGCTGAGGATGACCTTCTGTGAGGCTGCGAGCGCCCGGACATCCAGCCTTAAGACC 1680
1621 CTCGACGCTGAGGATGACCTTCTGTGAGGCTGCGAGCGCCCGGACATCCAGCCTTAAGACC 1680
1681 CACCAAGCCAGAGCGCAGGATGCCACGCTGTCCAGGGGAAGGGCTGTGATCTTTC 1740
1681 CACCAAGCCAGAGCGCAGGATGCCACGCTGTCCAGGGGAAGGGCTGTGATCTTTC 1740
1741 CATCGCTGGGCTGTGATCTGATCTCCAGACGCTGCGAGGCTGCGGCTGCGGCGCC 1800
1741 CATCGCTGGGCTGTGATCTGATCTCCAGACGCTGCGAGGCTGCGGCTGCGGCGCC 1800
1801 GTCCACTGACCCGAGGAGGCGGCTTTGGGCTCAACGCTTCGGGAGACTCGCCTGTCTC 1860
1801 GTCCACTGACCCGAGGAGGCGGCTTTGGGCTCAACGCTTCGGGAGACTCGCCTGTCTC 1860
1861 CTCACTGCTGCGCTCATTTCCAGGGAGCCAGGCGCGCGCTGCGCTTGAAGGAGGCGCAC 1920
1861 CTCACTGCTGCGCTCATTTCCAGGGAGCCAGGCGCGCGCTGCGCTTGAAGGAGGCGCAC 1920
1921 TACATTGTGTGAGTATGGGAGGCGGCTGAGGAGTGTGAGACACGCGGAGGCTGTGACG 1980
1921 TACATTGTGTGAGTATGGGAGGCGGCTGAGGAGTGTGAGACACGCGGAGGCTGTGACG 1980
1981 GAGCTGAAGGCTGCGGAGAGGCGGCGCGCAGCTGACGCTGTGCTGTGCTGCTGCTGCTG 2040
1981 GAGCTGAAGGCTGCGGAGAGGCGGCGCGCAGCTGACGCTGTGCTGTGCTGCTGCTGCTG 2040
2041 TCTAGACTGCGGAGCTTGGGAGACCGCGCGCGCTGCTGTGCGGCTGCGGAGGCTTCTTA 2100
2041 TCTAGACTGCGGAGCTTGGGAGACCGCGCGCGCTGCTGTGCGGCTGCGGAGGCTTCTTA 2100
2101 AGGAGCAGAGGAGGAGCTGCTGACAGACCCCGGCTGACGCTGGGCGGCTGCGGCGCC 2160
2101 AGGAGCAGAGGAGGAGCTGCTGACAGACCCCGGCTGACGCTGGGCGGCTGCGGCGCC 2160
2161 CTCCTCACTGAGGCTGCGGAGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 2220
2161 CTCCTCACTGAGGCTGCGGAGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 2220
2221 GCGGCACTGAGGAGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280
2221 GCGGCACTGAGGAGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280
2281 GGATCCCTGACGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
2281 GGATCCCTGACGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
2341 CCACCCAGAGGAGCTTCCGCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
2341 CCACCCAGAGGAGCTTCCGCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
2401 CCTGCTGCTGCTTAAAGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
2401 CCTGCTGCTGCTTAAAGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
2461 AAAAAAAAAA 2469
2461 AAAAAAAAAA 2469

RESULT 2
US-10-274-878-1
; Sequence 1, Application US/10274878

```

? Patent No. 6670163
?
? GENERAL INFORMATION:
?
? APPLICANT: RUSCH, Douglas et al
?
? TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
?
? TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
?
? TITLE OF INVENTION: THEREOF
?
? FILE REFERENCE: CL001099-CIP-DIV
?
? CURRENT APPLICATION NUMBER: US/10/274,878
?
? CURRENT FILING DATE: 2002-10-22
?
? PRIOR APPLICATION NUMBER: 09/849,334
?
? PRIOR FILING DATE: 2001-05-07
?
? PRIOR APPLICATION NUMBER: 09/773,371
?
? PRIOR FILING DATE: 2001-02-01
?
? NUMBER OF SEQ ID NOS: 4
?
? SOFTWARE: Fast-SEQ for Windows Version 4.0
?
? SEQ ID NO 1
?
? LENGTH: 2469
?
? TYPE: DNA
?
? ORGANISM: Human
?
? US-10-274-878-1

```

Qy	721	GCATGCGACCCCGACCGGAATGAGTCGGGCTCTGAGAGCTGCTCACAGACCTATTACAA	780
Db	721	GCATGCGACCCCGACCGGAATGAGTCGGGCTCTGAGAGCTCTCACAGACCTATTACAA	780
Qy	781	CAGCTGTGTCTCTGGATGCGGCTTCTCTACCCCTGCGAGAGCCTCGGGCTCTCTTC	840
Db	781	CAGCTGTGTCTCTGGATGCGGCTTCTCTACCCCTGCGAGAGCCTCGGGCTCTCTTC	840
Qy	841	CACCTGATCGACTCGCTTACTTGGGGTCCGGGCCAGCAGCGTCCCTGGGCTTCGAGAAG	900
Db	841	CACCTGATCGACTCGCTTACTTGGGGTCCGGGCCAGCAGCGTCCCTGGGCTTCGAGAAG	900
Qy	901	GGAGGGTCTCTTCAACATCGGTGCCCTCCACACGACAGATTGGGGCGCGCAGACCGC	960
Db	901	GGAGGGTCTCTTCAACATCGGTGCCCTCCACACGACAGATTGGGGCGCGCAGACCGC	960
Qy	961	TCCTGCACCGAGGTGCGCGCGCTATGAGAGCCTTCCAGAGGCGCTGAGGGGCTTC	1020
Db	961	TCCTGCACCGAGGTGCGCGCGCTATGAGAGCCTTCCAGAGGCGCTGAGGGGCTTC	1020
Qy	1021	AGCCTCTCTGAGGGAACTTCTCCATGCGGCCGAGCCAGACATGAGCGCTGCCTCCCTC	1080
Db	1021	AGCCTCTCTGAGGGAACTTCTCCATGCGGCCGAGCCAGACATGAGCGCTGCCTCCCTC	1080
Qy	1081	TGCGCACTGAGAGCTCATGATGGCCAGGCCAGGAATGTGTGTTGAGGGCTCTCA	1140
Db	1081	TGCGCACTGAGAGCTCATGATGGCCAGGCCAGGAATGTGTGTTGAGGGCTCTCA	1140
Qy	1141	CCACCTGCTCCATGGCCCCCAGACTGCTGCGGCCAGCTGCCTGCGCAGGAGGCC	1200
Db	1141	CCACCTGCTCCATGGCCCCCAGACTGCTGCGGCCAGCTGCCTGCGCAGGAGGCC	1200
Qy	1201	GCCAGGTGGCAGCCGAGTACAGGTAGTGCACCGACCATGGGCCAGCCACCGCTCCAC	1260
Db	1201	GCCAGGTGGCAGCCGAGTACAGGTAGTGCACCGACCATGGGCCAGCCACCGCTCCAC	1260
Qy	1261	GACTAGTGCCTGTCTCTGACTGCCCTGGTGCAATGTCGAAGGCCGAGTACTTCCGCTCC	1320
Db	1261	GACTAGTGCCTGTCTCTGACTGCCCTGGTGCAATGTCGAAGGCCGAGTACTTCCGCTCC	1320
Qy	1321	CTGGCCCACTACAGTAGCAGTGGCCCTCTGGACGGCTCCGAGCGACCGAGGAGAG	1380
Db	1321	CTGGCCCACTACAGTAGCAGTGGCCCTCTGGACGGCTCCGAGCGACCGAGGAGAG	1380
Qy	1381	CTCCCCAGCACGACGAGTCTTCTTGACGCCCCACCTCTTAAGCCCCCAGAGCCCT	1440
Db	1381	CTCCCCAGCACGACGAGTCTTCTTGACGCCCCACCTCTTAAGCCCCCAGAGCCCT	1440
Qy	1441	GTGCTGCGCAGAGTGGAGCGCAGCGAGCTTGGCAAGCACACCTGAAGCGTGCC	1500
Db	1441	GTGCTGCGCAGAGTGGAGCGCAGCGAGCTTGGCAAGCACACCTGAAGCGTGCC	1500
Qy	1501	ATCTGGGGCAGGAGGCGCTCGGCTGCAAGCCCTGTGCGCGCTCTGCGGAGGTG	1560
Db	1501	ATCTGGGGCAGGAGGCGCTCGGCTGCAAGCCCTGTGCGCGCTCTGCGGAGGTG	1560
Qy	1561	GACCTGTTTGGGCTGTGATCTCCACAGCCTCAGCGCTCACTGGCCAAAGTATGCGAG	1620
Db	1561	GACCTGTTTGGGCTGTGATCTCCACAGCCTCAGCGCTCACTGGCCAAAGTATGCGAG	1620
Qy	1621	CTGACCGTGAGATGATCTTGTGAGGCTGCGAGGCCCGGACATCCAGCCTAAGACC	1680
Db	1621	CTGACCGTGAGATGATCTTGTGAGGCTGCGAGGCCCGGACATCCAGCCTAAGACC	1680
Qy	1681	CACCAGAAGCCAGAGCCAGGATGCCACGCTGTCCAGGGGAGGGGCTGACATCTTC	1740
Db	1681	CACCAGAAGCCAGAGCCAGGATGCCACGCTGTCCAGGGGAGGGGCTGACATCTTC	1740
Qy	1741	CATCGGCTGGGCCCCCTGTCTGTGTTCTCAGCCAAAGACCGGTGGCGCTGTGGGGCC	1800
Db	1741	CATCGGCTGGGCCCCCTGTCTGTGTTCTCAGCCAAAGACCGGTGGCGCTGTGGGGCC	1800
Qy	1801	GTCACCTGACCCGAGGAGGGCGCTTTGGCTTCACGCTTCGGGGAGCTCGCTGTTC	1860

Db	1801	GTCCACTGACCCGAGGAGGGCGCTTTGGCCTCA	CGCTTCAGCGCTGGGAGACTCGCGCTGTC	1860
Qy	1861	CTCATCGTCGCGTCATTTCCAGGAGCCAGGCGCGCGCGCTGCGCTG	AGGAGGCGGAC	1920
Db	1861	CTCATCGTCGCGTCATTTCCAGGAGCCAGGCGCGCGCGCTGCGCTG	AGGAGGCGGAC	1920
Qy	1921	TACATTTGTGTCACTGAATGGGACCATGCAGGTG	TGTGAGACACGCGGAGGTGTGTGACG	1980
Db	1921	TACATTTGTGTCACTGAATGGGACCATGCAGGTG	TGTGAGACACGCGGAGGTGTGTGACG	1980
Qy	1981	GAGCTGAAGGCTCGGGAGAGGCGGCGCCAGCCT	TGCGAGTGTGTGCTGTGCTGCCAGC	2040
Db	1981	GAGCTGAAGGCTCGGGAGAGGCGGCGCCAGCCT	TGCGAGTGTGTGCTGTGCTGCCAGC	2040
Qy	2041	TCTTAGACTGCCAGCTTTGGGGHACCGCGGCCGT	CCTGCTGGGCCCCCAGGGGGCTTCTA	2100
Db	2041	TCTTAGACTGCCAGCTTTGGGGHACCGCGGCCGT	CCTGCTGGGCCCCCAGGGGGCTTCTA	2100
Qy	2101	AGGAGCCAGAGGAGGAGTGTGTTCAAGACCCCGGCAT	TCCAGTGGGCCAGTCCCCCGGCC	2160
Db	2101	AGGAGCCAGAGGAGGAGTGTGTTCAAGACCCCGGCAT	TCCAGTGGGCCAGTCCCCCGGCC	2160
Qy	2161	CTCTCTAACTGGHAGCCGAAAGGCCACGAGCGGCA	GAGCTGGAGGCTGCCCCAGCCCTGT	2220
Db	2161	CTCTCTAACTGGHAGCCGAAAGGCCACGAGCGGCA	GAGCTGGAGGCTGCCCCAGCCCTGT	2220
Qy	2221	GCCCCAGTGAAGCCAGCTCCGCGCTCATCTTGAAG	CACCCAGGCTGGCCGTGAGGGCCA	2280
Db	2221	GCCCCAGTGAAGCCAGCTCCGCGCTCATCTTGAAG	CACCCAGGCTGGCCGTGAGGGCCA	2280
Qy	2281	GGATCCCTGCA	CGCCCTCAGCGTGGCAAGCA	2340
Db	2281	GGATCCCTGCA	CGCCCTCAGCGTGGCAAGCA	2340
Qy	2341	CCACCCAGAGGACCTCCGGGCAATGCTGTCCGCGCT	CATGCTGGAGGCTGCCTCGGGCA	2400
Db	2341	CCACCCAGAGGACCTCCGGGCAATGCTGTCCGCGCT	CATGCTGGAGGCTGCCTCGGGCA	2400
Qy	2401	CTTGCTTGCCTATTAAAGACTGGTTCAGACCTGT	CTGAAAAA	2460
Db	2401	CTTGCTTGCCTATTAAAGACTGGTTCAGACCTGT	CTGAAAAA	2460
Qy	2461	AAAAAAA	2469	
Db	2461	AAAAAAA	2469	

```

RESULT 3
US-09-833-381-1293
; Sequence 1293, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1293
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1155)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1293

```

Query Match
38.3%; Score 946.8; DB 4; Length 1155;

	Best Local Similarity	99.2%; Pred. No.	1.4e-180;	Matches	Conservative	6; Mismatches	7; Indels	Gaps	1;																																																																																																																														
QY	1477	GGCAAGGCACACTTGAAGCGTGCCATCTTCGSGCAGGAGGCGCTCGCGGTGCACGCC	1536	Dd	133	GSCAAGGCACAACCCTAAAGCGTGCCATCTTCGSGCAGGAGGCGCTCGCGGTGCACGCC	192	QY	1537	TCTGTCGCGGTCTCGCGAGGTGCAGCTGTTTCGGSCTGTGATCTCCAGACGCTGCAG	1596	Dd	193	CTGTGC CGGCTCTCGCGAGGTGCAGCTGCTTCGGGCTGTGATCTCCAGACGCTGCAG	252	QY	1597	CCTCACCTGGCCCAAAGTAATGCGGAGCTGCACCGCTGAGGATGACTTCTGTGAGGCTCCGAG	1656	Dd	253	CCCCAACCGCCNAGTATGCGGAGCTCGACCGTAGGATGACTTCTGTGAGGCTCCGAG	312	QY	1657	GCCC GCGCATCCAGCTTAAGACCA CCAAGAAGCCAGAGGCCAGAGTSCCAGACCTGTCC	1716	Dd	313	GCCCCGGA CATCAGCCTAAGACCCACCA AAGACCAGAGGCCAGGATGCCACGCTGTCC	372	QY	1717	CAGGGAGAGGGCCTGACATCTTCATTCGGCTTGGGCGCCCTCTCTGTGTCTTCAGCCAAG	1776	Dd	373	CAGGGAGAGGGCCTGACATCTTCATTCGGCTTGGGCGCCCTCTCTGTGTCTTCAGCCAAG	432	QY	1777	AACCGGTGGCGGTGGTGGGCGCGTGCACCTGACCCGAGGAGAGGGCGGCTTGGCCCTC	1836	Dd	433	AAACCGGTGGCGGTGGTGGGCGCGGTGCACCTGACCCGAGGAGAGGGCGGCTTGGCCCTC	492	QY	1837	ACGCTTCGSGGAGACTCGCTGTCTCATCTGCCTTCATTCCAGGGAGCCAGGCGCGG	1896	Dd	493	ACGCTTCGSGGAGACTCGCTGTCTCATCTGCCTTCATTCCAGGGAGCCAGGCGCGG	552	QY	1897	GCGGCTGGCTGAAGAGGCGGCACTACATGTGTGATGTAATGGGCGAGCATGCAGGTGG	1956	Dd	553	GCGGCTGGCTGAAGAGGCGGCACTACATGTGTGATGTAATGGGCGAGCATGCAGGTGG	612	QY	1957	TGGAGACACGCGAGGTGTGACGGAGCTGAAGGCTGCGGAGAGAGCGCGCGCCAGCCTG	2016	Dd	613	TGAGAGACACGCGAGGTGTGACGGAGCTGAAGGCTGCGGAGAGAGCGCGCGCCAGCCTG	672	QY	2017	CAGGTGGTGTGCTGTGCCACGCTCTAGACTGCCACGCTTGGGGGACCGCGGCGCCGTC	2076	Dd	673	CAGGTGGTGTGCTGTGCCACGCTCTAGACTGCCACGCTTGGGGGACCGCGGCGCCGTC	732	QY	2077	CTGCTGGGCCCCAGGGGGCTTCTAAGAGCCAGAGGAGCATGTTGTCAAGACCCCGGCA	2136	Dd	733	CTGCTGGGCCCCAGGGGGCTTCTAAGAGCCAGAGGAGCATGTTGTCAAGACCCCGGCA	792	QY	2137	TCCACGTGGGCCAGTCCC GGCCCTCTCTCAACTGGAGCCGAAGGCCAGCAGGGCAAG	2196	Dd	793	TCCACGTGGGCCAGTCCC GGCCCTCTCTCAACTGGAGCCGAAGGCCAGCAGGGCAAG	852	QY	2197	ACTGGAGGCTGCCCCAGCGCTGTGCCAGTCAAGCCAGCTCCGCGCTCATCTTTGAAG	2256	Dd	853	ACTGGAGGCTGCCCCAGCGCTGTGCCAGTCAAGCCAGCTCCGCGCTCATCTTTGAAG	912	QY	2257	CACCCAGG GTGGCCGTGAGGGCCAGGATCCCTGCACGCCCTTCAGCCCTGGCTCCAGCTGG	2316	Dd	913	CACCCAGG GTGGCCGTGAGGGCCAGGATCCCTGCACGCCCTTCAGCCCTGGCTCCAGCTGG	971	QY	2317	CAGGAAGCACCGAGCATGCCCTCCACCCAGAGGACCTCCGGGCAATGCTGTCCGCGC	2376	Dd	972	CAGCAAGCACCGAGCATGCCCTCCACCCAGAGGACCTCCGGGCAATGCTGTCCGCGC	1031	QY	2377	TCATGCTGGAGGCTGCTCGGGGCACTTGCTGC GCCATAAAGACTGGTGCAGACTGTCTG	2436	Dd	1032	TCATGCTGGAGGCTGCTCGGGCACTTGCTGC GCCATAAAGACTGGTGCAGACTGTCTG	1091	QY	2437	AAAAAAAAAAAA 2446		Dd	1092	AAAAAAAAAAAA 1101	

RESULT 4

```
US-09-849-334-3
; Sequence 3, Application US/09849334
; Patent No. 6500655
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001099-CIP
; CURRENT APPLICATION NUMBER: US/09/849,334
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 19025
; TYPE: DNA
; ORGANISM: Human
US-09-849-334-3

Query Match
Best Local Similarity 14.9%; Score 368; DB 4; Length 19025;
Matches 379; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2058 GGGGACCGCCCGCCGCTCTGCTGGGCCCCCAGGGGGCTTCTTAAGGAGCCAGAGGGAGCA 2117
Db 15811 GGGGACCGCCCGCCGCTCTGCTGGGCCCCCAGGGGGCTTCTTAAGGAGCCAGAGGGAGCA 15870
QY 2118 TGGTTGCAAGACCCCGGCATCCACGTGGGCCAGTCCCGGCCCTCTCAACTGGAGCCG 2177
Db 15871 TGGTTGCAAGACCCCGGCATCCACGTGGGCCAGTCCCGGCCCTCTCAACTGGAGCCG 15930
QY 2178 AAAGCCCCAGGAGGCAAGACTGGAGGCTGCCCCAGCCCTGTGCCCGAGTGAAGCCAGC 2237
Db 15931 AAAGCCCCAGGAGGCAAGACTGGAGGCTGCCCCAGCCCTGTGCCCGAGTGAAGCCAGC 15990
QY 2238 TCCGCCCTCATCTTTGAAGCACCCAGGGTGGCGGTGAGGGCCAGGATCCCTGCACGCCCT 2297
Db 15991 TCCGCCCTCATCTTTGAAGCACCCAGGGTGGCGGTGAGGGCCAGGATCCCTGCACG-CCT 16049
QY 2298 CAGCCCTGGCTCCAGCTGGCAGCAAGACCGAGCATGCCCTCCCAACCCAGAGGACCTCC 2357
Db 16050 CAGCCCTGGCTCCAGCTGGCAGCAAGACCGAGCATGCCCTCCCAACCCAGAGGACCTCC 16109
QY 2358 GGGCAATGCTGTCCCGCTCATGCTGGAGGCTGCTGGGCACTGCTGGGCACTGCTGCCATTAAA 2417
Db 16110 GGGCAATGCTGTCCCGCTCATGCTGGAGGCTGCTGGGCACTGCTGGGCACTGCTGCCATTAAA 16169
QY 2418 GACTGGTCAGACCTGTCTGA 2437
Db 16170 GACTGGTCAGACCTGTCTGA 16189

RESULT 5
US-10-274-878-3
; Sequence 3, Application US/10274878
; Patent No. 6670163
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001099-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,878
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 19025
; TYPE: DNA
```

```
; ORGANISM: Human
US-10-274-878-3

Query Match
Best Local Similarity 14.9%; Score 368; DB 4; Length 19025;
Matches 379; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2058 GGGGACCGCCCGCCGCTCTGCTGGGCCCCCAGGGGGCTTCTTAAGGAGCCAGAGGGAGCA 2117
Db 15811 GGGGACCGCCCGCCGCTCTGCTGGGCCCCCAGGGGGCTTCTTAAGGAGCCAGAGGGAGCA 15870
QY 2118 TGGTTGCAAGACCCCGGCATCCACGTGGGCCAGTCCCGGCCCTCTCAACTGGAGCCG 2177
Db 15871 TGGTTGCAAGACCCCGGCATCCACGTGGGCCAGTCCCGGCCCTCTCAACTGGAGCCG 15930
QY 2178 AAAGCCCCAGGAGGCAAGACTGGAGGCTGCCCCAGCCCTGTGCCCGAGTGAAGCCAGC 2237
Db 15931 AAAGCCCCAGGAGGCAAGACTGGAGGCTGCCCCAGCCCTGTGCCCGAGTGAAGCCAGC 15990
QY 2238 TCCGCCCTCATCTTTGAAGCACCCAGGGTGGCGGTGAGGGCCAGGATCCCTGCACGCCCT 2297
Db 15991 TCCGCCCTCATCTTTGAAGCACCCAGGGTGGCGGTGAGGGCCAGGATCCCTGCACG-CCT 16049
QY 2298 CAGCCCTGGCTCCAGCTGGCAGCAAGACCGAGCATGCCCTCCCAACCCAGAGGACCTCC 2357
Db 16050 CAGCCCTGGCTCCAGCTGGCAGCAAGACCGAGCATGCCCTCCCAACCCAGAGGACCTCC 16109
QY 2358 GGGCAATGCTGTCCCGCTCATGCTGGAGGCTGCTGGGCACTGCTGGGCACTGCTGCCATTAAA 2417
Db 16110 GGGCAATGCTGTCCCGCTCATGCTGGAGGCTGCTGGGCACTGCTGGGCACTGCTGCCATTAAA 16169
QY 2418 GACTGGTCAGACCTGTCTGA 2437
Db 16170 GACTGGTCAGACCTGTCTGA 16189

RESULT 6
US-09-849-334-1/c
; Sequence 1, Application US/09849334
; Patent No. 6500655
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001099-CIP
; CURRENT APPLICATION NUMBER: US/09/849,334
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Human
US-09-849-334-1

Query Match
Best Local Similarity 2.8%; Score 68; DB 4; Length 2469;
Matches 158; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 1446 GCCCAGGAGCTGGAGGAGCCAGGAGCTTGGCAAGGCACACCTGAAGCGTGGCCATCCT 1505
Db 1753 GCCCAGGAGCTGGAGGAGCTTGGCAAGGCACACCTGAAGCGTGGCCATCCTGGCCT 1694
QY 1506 GGGGAGGAGGAGGCTGGCGCTGACGCCCTGACGCCCTGTCGCCGAGGTGGACCT 1565
Db 1693 CTGGCTTCTGGTGGGTCTTAGGCTGGATGTCGGGGCTCGGCAGCTCTCAGAGTCTAT 1634
QY 1566 GCTTCGGGCTGTGATCTCCCAAGAGCTGACGCTGACGCTGACGCTGACGCTGACGCTG 1625
Db 1633 CCTCAGGTCAGCTCCGCATCTTGGCCAGTGGAGCGCTGAGCGCTGAGCGATACAG 1574
QY 1626 CCGTGAGGATGACTTCTGTGAGGCTGCGGAGGCCGCCGACATCCAGCCTAAGACCCACCA 1685
```

```
Db      1573  CCAGAGCGGTCCACCTCGCAGGACGCGGCACAGGGGTGCAGCGCGCCCTCCT 1514
QY      1686  GAAGCCAGAGCGCCAGGATGCCAGCCCTGTCCAGGGGAAGGGCCCTGACATCTTCATCG 1745
Db      1513  CQTGCCCCAGGATGGCAGCGTTCAGGTGTGCTTTGCCAAGCTGCCTGGCTCCTCCAGCT 1454
QY      1746  GCTGGGGC 1753
Db      1453  CTTGGGCG 1446
```

RESULT 7

```
US-10-274-878-1/c
; Sequence 1, Application US/10274878
; Patent No. 6670163
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; FILE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; CURRENT APPLICATION NUMBER: US/10/274,878
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07/773,371
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Human
US-10-274-878-1
```

```
Query Match      2.8%; Score 68; DB 4; Length 2469;
Best Local Similarity 51.3%; Pred. No. 8.8e-05;
Matches 158; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY      1446  GCCGACGAGCTGGAGGAGCGCAGCGCTTGGCAAGGCACACCTGAAGCGTGCCATCCT 1505
Db      1753  GCCCAGCCGATGAAGATGTCAAGCCCTTCCCTGGGACAGCGGTGGCATCTCTGGCCT 1694
QY      1506  GGGCAGAGAGAGCGCTGGCGCTGACGCCCTGTGCGCGCTCTGCGGAGGTGGACCT 1565
Db      1693  CTGGCTTCTGTGGGTCTTAGGCTGGATGTCCGGGGCTCGGCAGCCTCACAGAAGTCAT 1634
QY      1566  GCTTCGGGCTGTGATCTCCAGACGCTGCAGCGCTCACTGGCCCAAGTATGCGAGCTCGA 1625
Db      1633  CTTACCGTTCAGCTCCGCTATCTTGGCCAGTGAGCGCTGAGCGCTCTGGAGATCACAG 1574
QY      1626  CCGTGAGGATGACTTCTGTGAGGCTGCGAGGCCCGGACATCCAGCTTAAGACCCACCA 1685
Db      1573  CCCGAAGCAGTCCACCTCGCGCAGGACGCGGCAAGCGGCTGCGAGCGCGCCCTCCT 1514
QY      1686  GAAGCCAGAGCCAGGATGCCAGCGCTGTCCAGGGGAAGGGCCCTGACATCTTCATCG 1745
Db      1513  CTTGCCCCAGGATGGCAGCGTTCAGGTGTGCTTTGCCAAGCTGCCTGGCTCCTCCAGCT 1454
QY      1746  GCTGGGGC 1753
Db      1453  CTTGGGCG 1446
```

RESULT 8

```
US-09-922-445-1
; Sequence 1, Application US/09922445
; Patent No. 6528268
; GENERAL INFORMATION:
; APPLICANT: Andersson, Maria K.
; APPLICANT: Berglund, Lars G. T.
; APPLICANT: Reneland, Rikard H.
```

```
; APPLICANT: Adam, Gail I. R.
; TITLE OF INVENTION: REAGENTS AND METHODS FOR DETECTION OF HEART FAILURE
; FILE REFERENCE: G6126US
; CURRENT APPLICATION NUMBER: US/09/922,445
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 38653
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(26156)
; OTHER INFORMATION:
; NAME/KEY: misc_feature
; LOCATION: (24801)..(24801)
; OTHER INFORMATION: nucleotide 24801 is a single nucleotide polymorphism which can be
; OTHER INFORMATION: A or G
; NAME/KEY: misc_feature
; LOCATION: (24941)..(24941)
; OTHER INFORMATION: nucleotide 24941 is a single nucleotide polymorphism which can be
; OTHER INFORMATION: T or C
; NAME/KEY: exon
; LOCATION: (26157)..(26252)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (26253)..(26401)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (26402)..(26543)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (26544)..(27024)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (27025)..(27178)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (27179)..(30519)
; OTHER INFORMATION:
; NAME/KEY: misc_feature
; LOCATION: (27645)..(27645)
; OTHER INFORMATION: nucleotide 27645 is a single nucleotide polymorphism which can be
; OTHER INFORMATION: C or G
; NAME/KEY: exon
; LOCATION: (30520)..(30681)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (30682)..(30894)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (30895)..(31027)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (31028)..(31747)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (31748)..(31841)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (31842)..(32400)
; OTHER INFORMATION:
; NAME/KEY: misc_feature
; LOCATION: (32163)..(32163)
; OTHER INFORMATION: nucleotide 32163 is a single nucleotide polymorphism which can be
; OTHER INFORMATION: A or C
; NAME/KEY: exon
; LOCATION: (32401)..(32528)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (32529)..(33414)
; OTHER INFORMATION:
```


QY 1553 GCGAGTGGACCTGCTTGGGCTGTGATCTCCAGACGCTGAGCGCTCACTGGCCAAAGT 1612
DB 22271 TCGCCGAGCACTGCGCCAGGCTTGGCCGAGACCGCGCGCTTGGGGGCTCTCTCGC 22330
QY 1613 ATCGGAGCTCGACCGCTGAGGATGACTTCTGTGAGGCTGCCGAGGCCCGCGGACATC 1668
DB 22331 TCTCGCCCTCGACGAAAGTCCCTCGCCGACCATGCGCGCTGCCGCGGACTC 22386

RESULT 10

US-08-456-837-6
; Sequence 6, Application US/08456837
; Patent No. 5643774
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,837
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/457,205
; FILING DATE: 01-JUN-1995
; APPLICATION NUMBER: 08/258,261
; FILING DATE: 08-Jun-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CSC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO

US-08-456-837-6
Query Match 2.3%; Score 57.2; DB 1; Length 28958;
Best Local Similarity 44.8%; Pred. No. 0.023;
Matches 267; Conservative 323; Indels 6; Gaps 1;

QY 1073 COTCCCTTGGCACTGAGCAGCTCATGATGCGCCGAGGCGCCAGGAATGTGTTGAGG 1132
DB 21977 CGACGATCTCGCCCGCGGAGCGGCTGCCCTGCCACCTTACCCCTTCAGCGCGGAGC 21856

QY 1133 GCCTCTCACCACTGCTCCATGCGCCCGCCCAAGACTGCTGGGCCAGCTGCGCTGGCGC 1192
DB 21857 GCTTCTGGCTCGAGCGCTTCCACGGCACCGCGCGCGTCAACACCTTGTCTCCGCTCG 21916
QY 1193 AGAGGCGCCCGAGTGGGAGCGGAGTACAGGCTAGTGACCGAGCAGCATGGCCAGCCAC 1252
DB 21917 AGGGGCGGTTCTTGGCAGGCGCATCGAGAGCGGGAATATCGACGCGCTCAGGGCGAGTCC 21976
QY 1253 CGTCCACGACTACGTGCTGTCTCTGACTGCCCTGGTGCATGTCAAGGGCCGAGTACT 1312
DB 21977 ACGTGACGGGACGAGAGCGCGCGCTTGGCTTGCCTTCCCTCCACACCTCGGAGCT 22036
QY 1313 TCCGCTTCCCTGGCCCACTACCACTAGCCATAGCCCTCTTGGGACGGGTCCCCAGGACCG 1372
DB 22037 TTGCGCCACGAGCGGCAAGAGCAGGCGACGCTCGACGCTTGGCGCTACCGCATCACCTGGA 22096
QY 1373 AGGAGAGCTCCCGACGACGAGCAGGTCTTCTGTCAGCCCCCGCCACTCTCTTAAGCCCC 1432
DB 22097 AGCCTCTGACCAACCGCCACCGCGCGCTGAGCGCGCTGAGCGAGGCGCCACCTGAGCCGAGGCTC 22270
QY 1433 GAGGCGCTGTGCTGCGCAGGAGCTGAGGAGCGGCGAGGCTTGGCAAGGCAACACTGA 1492
DB 22151 TCGTGGCGCGCTCTGGACGACGACGCGCTCCCTCCGCGCTCACCGAGGCGCTCGCC 22210
QY 1493 AGCCTGCCATCTTGGGCGCAGGAGGCGGTGCGGCTGTCAGCGCTTGTGCGCGGTCTCTGC 1552
DB 22211 GCGCGCGCGCGCTCTCGCGCTGCGCTGAGCGAGGCGCCACCTGAGCCGCGAGGCTC 22270
QY 1553 GCGAGTGGACCTGCTTGGGCTGTGATCTCCAGAGCTGCGAGCGCTCACTGGCCAAAGT 1612
DB 22271 TCGCCGAGCACTTGGCGCAGGCTTGGCGCGCGCTGCGCGCGCTCTCTCTCGC 22330
QY 1613 ATCGGAGCTCGACCGCTGAGGATGACTTCTGTGAGGCTGCGGAGGCGCCCGGACATC 1668
DB 22331 TCCTCGCCCTCGACGAAAGTCCCTCGCCGACCATGCGCGCTGCCCGCGGAGCTC 22386

RESULT 11

US-08-457-342-6
; Sequence 6, Application US/08457342
; Patent No. 5662898
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,342
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,205
; FILING DATE: 01-JUN-1995

```

; APPLICATION NUMBER: 08/259,261
; FILING DATE: 08-Jun-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-457-342-6

Query Match      2.3%; Score 57.2; DB 1; Length 28958;
Best Local Similarity 44.8%; Pred. No. 0.023;
Matches 267; Conservative 0; Mismatches 323; Indels 6; Gaps 1;

QY 1073 CGTCCCTCTGCGCACTGGAGAGCTCATGATGGCCAGCCAGGAGGAATGTGTTGAGG 1132
Db 21797 CGACGATCTGCGCCCGGAGCGGTGCCGTGCCACCTACCCCTTCCAGCGCGAGC 21856

QY 1133 GCCTCTCACCACTGCTCCATGGCCCCCAAGACTGCTGGCCAGCTGCGCTGGCGC 1192
Db 21857 GCTTCTGGCTCGACGCTCCAGCGCACCCGCGCGGGGTCAACACCTTCTCGCTCG 21916

QY 1193 AGGAGCGCCCGAGTGGAGTACAGGCTAGTGCACCGGACCATGCGCCAGGACT 1252
Db 21917 AGGGCGGTTCTGGCAGGCCATCGAGAGCGGAATATCGACGCTGCGCCCTTCCAGCGCGAGCTCC 21976

QY 1253 CCGTCTCCAGCACTGCTGCTCTCTGAGTGCCTGTGTGATGTCAGGCGGAGTACT 1312
Db 21977 AGTGGACGGCGACGAGACGCGCGCCCTTGCCCTTGCTCTTCCACCTTCGCGAGCT 22036

QY 1313 TCCGCTCCCTGGCCCACTACCACTGAGCATGGCCCTCTGCAACGCGCTCCCGAGCGCG 1372
Db 22037 TTGCGCCAGCGGCAAGAGCAGGCGACGGTTCGACGCTGGCGTACCGCATCATGTTGA 22096

QY 1373 AGGAGAGCTCCCGACGACGAGAGGTCTTCTGAGCGCCCGCCACCTCTTAAGCCCC 1432
Db 22097 AGCTCTGACCAACCGCCACACGCGCGCGAGCTGG-----CGGCGACCTGGCTCTCG 22150

QY 1433 GAGGCGCTGTGCTGCGCGAGGAGCTGGAGGCGCGAGGCGAGCTTGGCAGGCGACACCTGA 1492
Db 22151 TCGTCCCGCGCTCTGACGACGAGCGCTCCCTTCCGCGCTACCGAGGCGCTCGCCC 22210

QY 1493 AGCGTGCATCTGCGGCGAGGAGGCGCTGCGCTGCAAGCCCTGTGCGCGCTCTGCG 1552
Db 22211 GCGCGGCGCGCGCTCTGCGCGTGGCGCTGAGCCAGGCGCCACTTGGACCGGAGGCTC 22270

QY 1553 GCGAGTGGACTGCTTGGGCTGTGATCTCCAGACGCTGAGGCTCACTGCGCCAAAGT 1612
Db 22271 TCGCGGAGCACTGCGCGAGGCTTGGCGCGAGACCGCGCGCGCTGCTCTGCG 22330

QY 1613 ATCGGAGCTGACCGTGGAGTACTTCTGTGAGGCTGCGGAGCGCGCGGACATC 1668
Db 22331 TCCTGCGCTCGACAAAGTCCCTTGGCGGACCATGCGCGCTGCGCGGAGCTC 22386
```

RESULT 12

```

US-08-457-646A-6
; Sequence 6, Application US/08457646A
; Patent No. 5679560
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
```

```

; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hamner, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,646A
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,205
; FILING DATE: 01-JUN-1995
; APPLICATION NUMBER: 08/258,261
; FILING DATE: 08-Jun-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-457-646A-6
```

```

Query Match      2.3%; Score 57.2; DB 1; Length 28958;
Best Local Similarity 44.8%; Pred. No. 0.023;
Matches 267; Conservative 0; Mismatches 323; Indels 6; Gaps 1;

QY 1073 CGTCCCTCTGCGCACTGGAGAGCTCATGATGGCCAGCCAGGAGGAATGTGTTGAGG 1132
Db 21797 CGACGATCTGCGCCCGGAGCGGTGCCGTGCCACCTACCCCTTCCAGCGCGAGC 21856

QY 1133 GCCTCTCACCACTGCTCCATGGCCCCCAAGACTGCTGGCCAGCTGCGCTGGCGC 1192
Db 21857 GCTTCTGGCTCGACGCTCCAGCGCACCCGCGCGGGGTCAACACCTTCTCGCTCG 21916

QY 1193 AGGAGCGCCCGAGTGGAGTACAGGCTAGTGCACCGGACCATGCGCCAGGACT 1252
Db 21917 AGGGCGGTTCTGGCAGGCCATCGAGAGCGGAATATCGACGCTCAGCGCGAGCTCC 21976

QY 1253 CCGTCTCCAGCACTGCTGCTCTCTGAGTGCCTGTGTGATGTCAGGCGGAGTACT 1312
Db 21977 AGTGGACGGCGACGAGACGCGCGCCCTTGCCCTTGCTCTTCCACCTTCGCGAGCT 22036

QY 1313 TCCGCTCCCTGGCCCACTACCACTGAGCATGGCCCTCTGCAACGCGCTCCCGAGCGCG 1372
Db 22037 TTGCGCCAGCGGCAAGAGCAGGCGAGCTGGCGCTGCGCGTACCGCATCATGTTGA 22096

QY 1373 AGGAGAGCTCCCGACGACGAGAGGTCTTCTGAGCGCCCGCCACCTCTTAAGCCCC 1432
Db 22097 AGCTCTGACCAACCGCCACACGCGCGCGAGCTGG-----CGGCGACCTGGCTCTCG 22150

QY 1433 GAGGCGCTGTGCTGCGCGAGGAGCTGGAGGCGCGAGGCGAGCTTGGCAGGCGACACCTGA 1492
Db 22151 TCGTCCCGCGCTCTGACGACGAGCGCTCCCTTCCGCGCTACCGAGGCGCTCGCCC 22210

QY 1493 AGCGTGCATCTGCGGCGAGGAGGCGCTGCGCTGCAAGCCCTGTGCGCGCTCTGCG 1552
Db 22211 GCGCGGCGCGCGCTCTGCGCGTGGCGCTGAGCCAGGCGCCACTTGGACCGGAGGCTC 22270

QY 1553 GCGAGTGGACTGCTTGGGCTGTGATCTCCAGACGCTGAGGCTCACTGCGCCAAAGT 1612
Db 22271 TCGCGGAGCACTGCGCGAGGCTTGGCGCGAGACCGCGCGCGCTGCTCTGCG 22330

QY 1613 ATCGGAGCTGACCGTGGAGTACTTCTGTGAGGCTGCGGAGCGCGGACATC 1668
Db 22331 TCCTGCGCTCGACAAAGTCCCTTGGCGGACCATGCGCGCTGCGCGGAGCTC 22386
```


Db 22097 AGCCTGTACACCGCCACCGCCGACCTGG-----CGGGACCTGTGCTCTCG 22150
QY 1433 GAGGCCCTGTGTCGCGCAGGAGCTGGAGGAGCGCAGCTTGGCAAGGCACACCTGA 1492
Db 22151 TCGTGGCGCGCTCTGGACGACGACGCTCCCTCCGCGCTCACCGAGCGCTGCGCC 22210
QY 1493 AGCGTGCCATCTCTGGGGCAGGAGCGCTTGGCGGTGCAAGCGCTGTGCGCGTCTGTC 1552
Db 22211 GGGCGCGCGCGCGCTCTCGCGCTGCGCTGAGCCAGGCGCCACCTGGACCGGAGGCTC 22270
QY 1553 GCGAGGTGACCTGCTTTCGGGCTGTGATCTCCAGAGCTGACGCTCACTGGCCAAAGT 1612
Db 22271 TCGCGGAGCACCTTGGCGGAGCTTGGCCGAGACCGCGCGCTCGCGGCGTGTCTCGC 22330
QY 1613 ATCGGAGCTCGACCGTGTGAGGATGACTTCTGTGAGGCTGCGGAGGCGCCGACATC 1668
Db 22331 TCCTGCGCCTCGACGAAAGTCCCTCGCGGACCATGCGCGCGTGGCGCGGAGCTC 22386

RESULT 13

US-08-458-076A-6
; Sequence 6, Application US/08458076A
; Patent No. 5698425
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/458,076A
; APPLICATION NUMBER: US/08/457,205
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,205
; FILING DATE: 01-JUN-1995
; APPLICATION NUMBER: 08/258,261
; FILING DATE: 08-Jun-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO

; ANTI-SENSE: NO
US-08-458-076A-6
Query Match 2.3%; Score 57.2; DB 1; Length 28958;
Best Local Similarity 44.8%; Pred. No. 0.023;
Matches 267; Conservative 0; Mismatches 323; Indels 6; Gaps 1;
QY 1073 CGTCCCTCTGCGACCTGGAGCAGCTCATGATGCCGCCAGAGGCCAGGAATGTGTGTTGAGG 1132
Db 21797 CGAGATTCCTGCCCCCGGGAAGCGGTGCGCTGCCACCTACCCCTTCCAGGGCGAGC 21856
QY 1133 GCCTCTCACACACTGCTCCATGCGCCCCCAAGACTGCTTGGCCCAAGCTGGGCTGGCGC 1192
Db 21857 GCTTCTGGCTCGACGCTCCACGCGCACCGCGCGCGGTCAACCACTTGTCTCGCTCG 21916
QY 1193 AGAGGCGCGGAGGTGGCAGCGAGTACAGGCTAGTGACCGGACCATGCCCCAGCCAC 1252
Db 21917 AGGGGCGGTTCTTGGCAGGCGCATCGAGCGGGAATATCGACGCGCTCAGCGGCGAGCTC 21976
QY 1253 CCGTCCACGACTAGTGCCTGTCTCTTGGACTGCCCCCTGTGTGTCATGTCAAGGCGGAGTACT 1312
Db 21977 AGTGGACGCGGAGGAGCAGCGCGCGCTTGGCCCTGCTCTTCCACCTCGCGAGCT 22036
QY 1313 TCCGCTCCCTGGCCCCACTACACGTAGCCATGCGCCCTTTCGAGGCTTCCGAGCGGCTCC 1372
Db 22037 TTGCGCACGAGCGCAAGAGCAGGCGCACGCTCGACGCTGCGCGCTACCGCATCACGTGA 22096
QY 1373 AGGAGAGCTCCCCACGCGCAGCAGCTTCTTCTGCGAGCCCCCAGCTTCTTAAGCCCC 1432
Db 22097 AGCTCTGACACCGCGCCACCGCCCGGCGGCTG-----CGGACCTTGGCTCTCG 22150
QY 1433 GAGGCCCTGTGCTGCGCAGGAGCTGGAGGAGCGCAGGAGCTTGGCAAGGCACACCTGA 1492
Db 22151 TCGTGGCGCGCTCTGGACGACGACGCTCCCTCCGCGCTCACGAGCGCTCGGCC 22210
QY 1493 AGCGTGCCATCTCTGGGGCAGGAGCGGCTGCGGCTGCGAGCGCTGTGCGCGCTCTGTC 1552
Db 22211 GGGCGCGCGCGCGCTCTGCGCTGCGCTGAGCCAGGCGCCACCTGAGCGCGAGGCTC 22270
QY 1553 GCGAGGTGAGCTGCTTTCGGGCTGTGATCTCCAGAGCTGCGAGCGCTCACTGGCCAAAGT 1612
Db 22271 TCGCGGAGCACCTTGGCGCAGCTTGGCCGAGACCGCGCGCTCGCGGCGTGTCTCGC 22330
QY 1613 ATCGGAGCTCGACCGTGTGAGGATGACTTCTGTGAGGCTGCGAGGCGCCCGGACATC 1668
Db 22331 TCCTGCGCCTCGACGAAAGTCCCTCGCGGACCATGCGCGCGTGGCGCGGAGCTC 22386
RESULT 14
US-08-764-233A-4
; Sequence 4, Application US/08764233A
; Patent No. 5716849
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Schupp, Thomas
; APPLICANT: Beck, James J.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Neff, Snezana
; APPLICANT: Ryals, John A.
; TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/764,233A
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/729,214
;; FILING DATE: 09-OCT-1996
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/258,261
;; FILING DATE: 08-JUN-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Meigs, J. Timothy
;; REGISTRATION NUMBER: 38,241
;; REFERENCE/DOCKET NUMBER: 1506/CIP6
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (919) 541-8587
;; TELEFAX: (919) 541-8689
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 28958 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; ORIGINAL SOURCE:
;; ORGANISM: Sorangium cellulosum
;; IMMEDIATE SOURCE:
;; CLONE: p98/1
US-08-764-233A-4

Query Match 2.3%; Score 57.2; DB 1; Length 28958;
Best Local Similarity 44.8%; Pred. No. 0.023;
Matches 267; Conservative 0; Mismatches 323; Indels 6; Gaps 1;
QY 1073 CGTCTCTGCGCACTGAGCAGCTCATGATGCGCCAGGCCAGGAAATGTGTTGAGG 1132
Db 21797 CGACGATCTGCCCCCGGAGCGGTGCGCTGCCACCTACCCCTTCAGCGGAGC 21856
QY 1133 GCCTCTACCACTGCTTCATGCGCCCCCAAGACTGCTGCGCCAGCTGCGCTGAGG 1192
Db 21857 GCTTCTGCTGCGCACTGAGCAGCTGCGCGCGCGGTCAACACCTTGTCTCGCTCG 21916
QY 1193 AGGAGCGCGCCAGGTGCGCAGCCAGTACAGGCTAGTGACCGGACCATGCGCCAGCAC 1252
Db 21917 AGGGCGGTCTGGCAGGCGCATGAGAGCGGGAATATCGACCGCTCAGCGCGCAGCTCC 21976
QY 1253 CGTCCAGCACTACGTGCTGTCTCTGACTGCGCTGTGTGATGTCAAGGCGGAGTACT 1312
Db 21977 AGTGGACGGGACGAGCAGCGCGCGCTTGCCTGTCTCTTCCACCTTCGCGAGCT 22036
QY 1313 TCGCTCTCGTGGCCACTACCAAGTAGGCTAGCGCTTGGCAAGGCTTCCCGAGCGAG 1372
Db 22037 TTGCGCACGAGCGGCAAGCAGGCGCGGTGCGCGCTGCGGCTACCGCATCATGCTGA 22096
QY 1373 AGGAGAGCTCCCAAGCAGCAGCAGGTCTTCTGCGCCCCCAGCTCTTAAGCCCC 1432
Db 22097 AGCTCTGACACACCGCACCGCGCGCGCTG-----CCGGCACCTGGCTCTCG 22150
QY 1433 GAGGCCCTGTGTGCGCGCAGGAGCTGGAGGCGCGAGCTGGCAAGGCACACCTGA 1492
Db 22151 TGTGCGCGCGCTCTGGACGAGCAGCGCTCTTCCGCGCTTACAGGCGCTTCGCGC 22210
QY 1493 AGCGTGCCATCTTGGGCGAGGAGGCGCTGCGCTGCAAGCCCTGTGCGCGCTCTTC 1552
Db 22211 GCGCGGCGCGCGCTCTCTCGCGTGGCGCTGAGCCAGCCCACTGGACCGCGAGGCTC 22270
QY 1553 GCGAGGTGCACTGCTTGGGCTGTGATCTCCACAGCTGCGAGGCTTCTATGCGCAAGT 1612
Db 22271 TCGCGAGACCTGCGCGAGGCTTGGCGGAGACCGCGCGCGCTGCGCGCTGCTCTCG 22330
QY 1613 ATCGGAGCTGACCGTGGAGGATGATCTTGTGAGGCTGCGAGGCCCGCGGACATC 1668
Db 22331 TCCTCGCCTCGACGAAAGTCCCTCGCGGACCATGCGCGCTGCGCGCGGAGCTC 22386

RESULT 15

US-08-457-335A-6
; Sequence 6, Application US/08457335A
; Patent No. 5723759
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,335A
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,205
; FILING DATE: 01-JUN-1995
; APPLICATION NUMBER: 08/258,261
; FILING DATE: 08-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-457-335A-6

Query Match 2.3%; Score 57.2; DB 1; Length 28958;
Best Local Similarity 44.8%; Pred. No. 0.023;
Matches 267; Conservative 0; Mismatches 323; Indels 6; Gaps 1;
QY 1073 CGTCTCTGCGCACTGAGCAGCTCATGATGCGCCAGGCCAGGAAATGTGTTGAGG 1132
Db 21797 CGACGATCTGCCCCCGGAGCGGTGCGCTGCCACCTACCCCTTCAGCGGAGC 21856
QY 1133 GCCTCTACCACTGCTTCATGCGCCCCCAAGACTGCTGCGCCAGCTGCGCTGAGG 1192
Db 21857 GCTTCTGCTGCGCACTGAGCAGCTGCGCGCGGTCAACACCTTGTCTCGCTCG 21916
QY 1193 AGGAGCGCGCCAGGTGCGCAGCCAGTACAGGCTAGTGACCGGACCATGCGCCAGCAC 1252
Db 21917 AGGGCGGTCTGGCAGGCGCATGAGAGCGGGAATATCGACCGCTCAGCGCGCAGCTCC 21976

Tue Jul 13 15:19:49 2004

[illegible]

Search completed: July 13, 2004, 08:26:32
Job time : 129.439 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2004, 12:07:02 ; Search time 427 Seconds
(without alignments)
528.143 Million cell updates/sec

Title: US-10-697-266-2

Perfect score: 3774

Sequence: 1 MILERRPDGAGGESPRLO.....QPCAPVKPAPPSLKHPGWP 723

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1279676 seqs, 311918243 residues

Total number of hits satisfying chosen parameters: 1279676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3774	100.0	723	12	US-10-697-266-2
2	3774	100.0	723	14	US-10-274-878-2
3	3449.5	91.4	670	12	US-10-451-207-2
4	3435	91.0	669	12	US-10-113-944-318
5	2331	61.8	634	12	US-10-697-266-4
6	2331	61.8	634	14	US-10-274-878-4
7	2331	61.8	643	15	US-10-239-607-71
8	1442.5	38.2	311	16	US-10-408-765A-2465
9	1278	33.9	257	9	US-09-925-301-1016
10	1278	33.9	257	14	US-10-023-896-59
11	1278	33.9	257	14	US-10-023-896-87
12	1256	33.3	686	15	US-10-108-260A-4840
13	1255	33.3	611	12	US-10-092-900A-352
14	1255	33.3	686	15	US-10-239-607-53
15	1249	33.1	686	9	US-09-895-040A-3

16	1242	32.9	686	15	US-10-239-607-73	Sequence 73, Appl
17	1228	32.5	686	16	US-10-408-765A-2644	Sequence 2644, Ap
18	1195	31.7	686	12	US-10-092-900A-350	Sequence 350, App
19	1023	27.1	718	15	US-10-239-607-72	Sequence 72, Appl
20	936	24.8	180	12	US-10-112-944-776	Sequence 776, App
21	849	22.5	351	15	US-10-378-029-101	Sequence 101, App
22	394	10.4	298	15	US-10-108-260A-2912	Sequence 2912, Ap
23	387	10.3	1403	12	US-10-425-114-54518	Sequence 54518, A
24	380.5	10.1	939	15	US-10-094-749-1895	Sequence 1895, Ap
25	380.5	10.1	1636	16	US-10-433-794-1	Sequence 1, Appl
26	301	8.0	877	12	US-10-425-114-59396	Sequence 59396, A
27	294.5	7.8	883	16	US-10-437-963-180574	Sequence 180574, A
28	274.5	7.3	868	12	US-10-221-278-251	Sequence 251, App
29	274.5	7.3	868	15	US-10-291-172-251	Sequence 251, App
30	268	7.1	921	9	US-09-800-729-199	Sequence 199, App
31	268	7.1	921	15	US-10-264-049-3015	Sequence 3015, Ap
32	264.5	7.0	868	9	US-09-800-729-106	Sequence 106, App
33	264.5	7.0	868	11	US-09-833-245-2179	Sequence 2179, Ap
34	251.5	6.7	200	12	US-10-424-599-275853	Sequence 275853, A
35	251.5	6.7	397	16	US-10-437-963-123181	Sequence 123181, A
36	232	6.1	661	9	US-09-801-368-166	Sequence 166, App
37	185	4.9	104	15	US-10-630-590-187	Sequence 187, App
38	156.5	4.1	780	9	US-09-771-161A-144	Sequence 144, App
39	156.5	4.1	942	14	US-10-354-358-74	Sequence 74, Appl
40	155.5	4.1	942	9	US-09-771-161A-235	Sequence 235, App
41	155.5	4.1	942	9	US-09-771-161A-236	Sequence 236, App
42	146	3.9	1151	15	US-10-369-493-7865	Sequence 7865, Ap
43	145.5	3.9	1522	14	US-10-134-102-8	Sequence 8, Appl
44	145.5	3.9	1522	15	US-10-144-198-31	Sequence 31, Appl
45	144.5	3.8	77	12	US-10-042-865-86	Sequence 86, Appl

ALIGNMENTS

RESULT 1
US-10-697-266-2
; Sequence 2, Application US/10697266
; Publication No. US20040067522A1
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001099-CIP-DIV2
; CURRENT APPLICATION NUMBER: US/10/697,266
; CURRENT FILING DATE: 2003-10-31
; PRIOR FILING DATE: 2002-10-22
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-697-266-2

Query Match 100.0%; Score 3774; DB 12; Length 723;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MILERRPDGAGGESPRLOISRRPKTRVSSLRGRREGLRDVCAMRGCRVHRGEDPVR	60
Db	1	MILERRPDGAGGESPRLOISRRPKTRVSSLRGRREGLRDVCAMRGCRVHRGEDPVR	60
QY	61	VHVGPMPQLHAVGCDLSLTQCGQLQSRRAQIHQOIKELQMRGTGAENLYRATSNRVR	120
Db	61	VHVGPMPQLHAVGCDLSLTQCGQLQSRRAQIHQOIKELQMRGTGAENLYRATSNRVR	120

```
QY 121 ETVALELSYNSNQLLKEELELSGGVDPGRHSGEAVTVPMIPLGLKETKELDWSTPLK 180
DB 121 ETVALELSYNSNQLLKEELELSGGVDPGRHSGEAVTVPMIPLGLKETKELDWSTPLK 180
QY 181 ELISVHFEGDGASYEAREIREALRQAMRTPSRNESGLELLTAYYNQICFLDARFLTPAR 240
DB 181 ELISVHFEGDGASYEAREIREALRQAMRTPSRNESGLELLTAYYNQICFLDARFLTPAR 240
QY 241 SLGLFFHYDLSLTGVPAAQORALAFKESVLFNIGALHTQI GARODRSCTEGARAMEAFQ 300
DB 241 SLGLFFHYDLSLTGVPAAQORALAFKESVLFNIGALHTQI GARODRSCTEGARAMEAFQ 300
QY 301 RAAAGAFSLLENFNSHAPSPDMSAASLCALEQLMMAQAQECVFEGLSPASMAPQDCLAQL 360
DB 301 RAAAGAFSLLENFNSHAPSPDMSAASLCALEQLMMAQAQECVFEGLSPASMAPQDCLAQL 360
QY 361 RLAQEAQAQAAEYRLVHRTMAQPPVHDYVPVSWTALVHVKAIEYFRSLAHYHVAMALCDGS 420
DB 361 RLAQEAQAQAAEYRLVHRTMAQPPVHDYVPVSWTALVHVKAIEYFRSLAHYHVAMALCDGS 420
QY 421 PATEGELTHQVFLQPPFTSSKPRGPVLPOELEERRQLGKAHLKRAILGQEAALRLHALC 480
DB 421 PATEGELTHQVFLQPPFTSSKPRGPVLPOELEERRQLGKAHLKRAILGQEAALRLHALC 480
QY 481 RVLREVDLLRAVISQTLQRLSLAKYAELDRDDFCEAAEAPDIQPKTHQKPEARMPLRSQ 540
DB 481 RVLREVDLLRAVISQTLQRLSLAKYAELDRDDFCEAAEAPDIQPKTHQKPEARMPLRSQ 540
QY 541 KPDI FHLRGLSVFSAKNRWLVGPVHLTRGEGGFGTLRGDSVPLTAAVIPGSOAAA 600
DB 541 KPDI FHLRGLSVFSAKNRWLVGPVHLTRGEGGFGTLRGDSVPLTAAVIPGSOAAA 600
QY 601 GLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAEAGASLQVWSLLPSSRLPSLGDRRPVLL 660
DB 601 GLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAEAGASLQVWSLLPSSRLPSLGDRRPVLL 660
QY 661 GPRGLLRSOREHGCKTPASTWASPRLLNWSRKAQGGTGGCPQPCAPVKPAPPSLKH 720
DB 661 GPRGLLRSOREHGCKTPASTWASPRLLNWSRKAQGGTGGCPQPCAPVKPAPPSLKH 720
QY 721 GWP 723
DB 721 GWP 723
```

RESULT 2

```
US-10-274-878-2
; Sequence 2, Application US/10274878
; Publication NO. US20030049792A1
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THERSOF
; FILE REFERENCE: CL001099-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,878
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 2
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Human
US-10-274-878-2
```

Query Match 100.0%; Score 3774; DB 14; Length 723;
Best Local Similarity 100.0%; Pred No. 0;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MILEERPDGAGAGBESPRQLQISRRPKRTRVSSLRGRREGRLDYCAMRGCRVHRGEDPVR 60
DB 1 MILEERPDGAGAGBESPRQLQISRRPKRTRVSSLRGRREGRLDYCAMRGCRVHRGEDPVR 60
QY 61 VHVCPMPOLHANGCDSLTQICCGLOSRRAQIHQOQIDKELQMRITGAENLYRATSNNRVR 120
DB 61 VHVCPMPOLHANGCDSLTQICCGLOSRRAQIHQOQIDKELQMRITGAENLYRATSNNRVR 120
QY 121 ETVALELSYNSNQLLKEELELSGGVDPGRHSGEAVTVPMIPLGLKETKELDWSTPLK 180
DB 121 ETVALELSYNSNQLLKEELELSGGVDPGRHSGEAVTVPMIPLGLKETKELDWSTPLK 180
QY 181 ELISVHFEGDGASYEAREIREALRQAMRTPSRNESGLELLTAYYNQICFLDARFLTPAR 240
DB 181 ELISVHFEGDGASYEAREIREALRQAMRTPSRNESGLELLTAYYNQICFLDARFLTPAR 240
QY 241 SLGLFFHYDLSLTGVPAAQORALAFKESVLFNIGALHTQI GARODRSCTEGARAMEAFQ 300
DB 241 SLGLFFHYDLSLTGVPAAQORALAFKESVLFNIGALHTQI GARODRSCTEGARAMEAFQ 300
QY 301 RAAAGAFSLLENFNSHAPSPDMSAASLCALEQLMMAQAQECVFEGLSPASMAPQDCLAQL 360
DB 301 RAAAGAFSLLENFNSHAPSPDMSAASLCALEQLMMAQAQECVFEGLSPASMAPQDCLAQL 360
QY 361 RLAQEAQAQAAEYRLVHRTMAQPPVHDYVPVSWTALVHVKAIEYFRSLAHYHVAMALCDGS 420
DB 361 RLAQEAQAQAAEYRLVHRTMAQPPVHDYVPVSWTALVHVKAIEYFRSLAHYHVAMALCDGS 420
QY 421 PATEGELTHQVFLQPPFTSSKPRGPVLPOELEERRQLGKAHLKRAILGQEAALRLHALC 480
DB 421 PATEGELTHQVFLQPPFTSSKPRGPVLPOELEERRQLGKAHLKRAILGQEAALRLHALC 480
QY 481 RVLREVDLLRAVISQTLQRLSLAKYAELDRDDFCEAAEAPDIQPKTHQKPEARMPLRSQ 540
DB 481 RVLREVDLLRAVISQTLQRLSLAKYAELDRDDFCEAAEAPDIQPKTHQKPEARMPLRSQ 540
QY 541 KPDI FHLRGLSVFSAKNRWLVGPVHLTRGEGGFGTLRGDSVPLTAAVIPGSOAAA 600
DB 541 KPDI FHLRGLSVFSAKNRWLVGPVHLTRGEGGFGTLRGDSVPLTAAVIPGSOAAA 600
QY 601 GLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAEAGASLQVWSLLPSSRLPSLGDRRPVLL 660
DB 601 GLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAEAGASLQVWSLLPSSRLPSLGDRRPVLL 660
QY 661 GPRGLLRSOREHGCKTPASTWASPRLLNWSRKAQGGTGGCPQPCAPVKPAPPSLKH 720
DB 661 GPRGLLRSOREHGCKTPASTWASPRLLNWSRKAQGGTGGCPQPCAPVKPAPPSLKH 720
QY 721 GWP 723
DB 721 GWP 723
```

RESULT 3

```
US-10-451-207-2
; Sequence 2, Application US/10451207
; Publication NO. US20040038267A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: BURFORD, Neil
; APPLICANT: DING, Li
; APPLICANT: THORNTON, Michael B.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: SWARNAKAR, Anica
; APPLICANT: DUGGAN, Brendan M.
; APPLICANT: LU, Dying Aina M.
; APPLICANT: THANGAVELU, Kavitha
; APPLICANT: WARREN, Bridget A.
; APPLICANT: TANG, Y. Tom
```

APPLICANT: KHAN, Farrah A.
APPLICANT: YAO, Monique G.
APPLICANT: EVERLING, Brooke M.
TITLE OF INVENTION: INTRACELLULAR SIGNALING MOLECULES
FILE REFERENCE: PF-0868 USN
CURRENT FILING DATE: 2003-06-18
PRIOR APPLICATION NUMBER: US/10/451,207
CURRENT FILING DATE: 2003-06-18
PRIOR APPLICATION NUMBER: PCT/US01/50315
PRIOR FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: US 60/257,804
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: US 60/260,102
PRIOR FILING DATE: 2001-01-05
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PERL Program
SEQ ID NO 2
LENGTH: 670
TYPE: PRT
ORGANISM: Homo sapiens
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040038267A1 5560161CD1
US-10-451-207-2

Query Match 91.4%; Score 3449.5; DB 12; Length 670;
Best Local Similarity 92.7%; Pred. No. 1.9e-293;
Matches 670; Conservative 0; Mismatches 0; Indels 53; Gaps 1;
QY 1 MILERRPGAGAGEESPRIQISRRKPKRTXVSSLRGRREGRLDVCAMWRCRVHRGDDPVR 60
DB 1 MILERRPGAGAGEESPRIQ----- 20
QY 61 VHVGPMPQLHAGVCDLSLTQICGQLOSRAQIHQOIDKELQMTGAENLYRATSNRVR 120
DB 21 -----GCDLSLTQICGQLOSRAQIHQOIDKELQMTGAENLYRATSNRVR 67
QY 121 ETVALELYSVNSNLQILKEELELSGGVDPGRHSGEAVTVPMIPGLKTKELDWSLTK 180
DB 68 ETVALELYSVNSNLQILKEELELSGGVDPGRHSGEAVTVPMIPGLKTKELDWSLTK 127
QY 181 ELISVHFGEDGASYEAEIREALRQAMRTPSRNSGLELLTAYYNQLCFDARELTTPAR 240
DB 128 ELISVHFGEDGASYEAEIREALRQAMRTPSRNSGLELLTAYYNQLCFDARELTTPAR 187
QY 241 SGLGFPHWYDLSLTGVPDAQORALAFEGSVLFNIGALHTQIGARQDRSCTEGARRAMEAFQ 300
DB 188 SGLGFPHWYDLSLTGVPDAQORALAFEGSVLFNIGALHTQIGARQDRSCTEGARRAMEAFQ 247
QY 301 RAAGAFSLRENFSHAPSPDMGSAASLCALBQLMMAQAQECVPEGLSPAPMAPODCLAQL 360
DB 248 RAAGAFSLRENFSHAPSPDMGSAASLCALBQLMMAQAQECVPEGLSPAPMAPODCLAQL 307
QY 361 RLQAQAAQAAAYRLVHRTMAQPPVHDYVPVSWTALVHVKAIFYRSLAHYHVAMALCDGS 420
DB 308 RLQAQAAQAAAYRLVHRTMAQPPVHDYVPVSWTALVHVKAIFYRSLAHYHVAMALCDGS 367
QY 421 RATEGELPHEOVFLOPPSTSKPRGVLPQELERQOLKAHLKRAILQOEALRHLALC 480
DB 368 RATEGELPHEOVFLOPPSTSKPRGVLPQELERQOLKAHLKRAILQOEALRHLALC 427
QY 481 RVLREVDLLRAVISQTLQSLAKYAELEDREDDFCEAAEAPDIQPKTHQKPEARMPELSOG 540
DB 428 RVLREVDLLRAVISQTLQSLAKYAELEDREDDFCEAAEAPDIQPKTHQKPEARMPELSOG 487
QY 541 KGPDI FHLRGPI SVFSAXNRWRLVGPVHLTRGEGGFLTRGDSVPVLI AAVIPGSOAAA 600
DB 488 KGPDI FHLRGPI SVFSAXNRWRLVGPVHLTRGEGGFLTRGDSVPVLI AAVIPGSOAAA 547
QY 601 GLKEGDYIVSVNGQPCRRWRHAEVWTELKAAEAGASLQVSVLLPSSRLPSIGDRRPVLL 660
DB 548 GLKEGDYIVSVNGQPCRRWRHAEVWTELKAAEAGASLQVSVLLPSSRLPSIGDRRPVLL 607
QY 661 GPRGLLRSORRHGCKTPASTWASPRLLNWSRKAQOQKGTGGCPQPCAPVKPAPPSSLKHP 720

DB 608 GPRGLLRSORRHGCKTPASTWASPRLLNWSRKAQOQKGTGGCPQPCAPVKPAPPSSLKHP 667
QY 721 GWP 723
DB 668 GWP 670
RESULT 4
US-10-112-944-318
Sequence 318, Application US/10112944
Publication No. US20040048249A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Yang, Yonghong
APPLICANT: Weng, Gezhi
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Xue, Aidong J.
APPLICANT: Wang, Jian-Rui
APPLICANT: Wehrman, Tom
APPLICANT: Ghosh, Malabika
APPLICANT: Wang, Dunrui
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Zhiwei
TITLE OF INVENTION: No. US20040048249A1 Secreted Polypeptides
TITLE OF INVENTION: 805A
FILE REFERENCE: 805A
CURRENT APPLICATION NUMBER: US/10/112,944
PRIOR FILING DATE: 2002-03-28
PRIOR APPLICATION NUMBER: US 09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US 09/491,404
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: US 09/496,914
PRIOR FILING DATE: 2000-02-03
PRIOR APPLICATION NUMBER: US 09/515,126
PRIOR FILING DATE: 2000-02-28
PRIOR APPLICATION NUMBER: US 09/519,705
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: US 09/552,929
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: US 09/577,408
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 924
SOFTWARE: pt_FL_genes Version 5.0
SEQ ID NO 318
LENGTH: 669
TYPE: PRT
ORGANISM: Homo sapiens
US-10-112-944-318
Query Match 91.0%; Score 3435; DB 12; Length 669;
Best Local Similarity 92.5%; Pred. No. 3.6e-292;
Matches 669; Conservative 0; Mismatches 0; Indels 54; Gaps 2;
QY 1 MILERRPGAGAGEESPRIQISRRKPKRTXVSSLRGRREGRLDVCAMWRCRVHRGDDPVR 60
DB 1 MILERRPGAGAGEESPRIQ----- 20
QY 61 VHVGPMPQLHAGVCDLSLTQICGQLOSRAQIHQOIDKELQMTGAENLYRATSNRVR 120
DB 21 -----GCDLSLTQICGQLOSRAQIHQOIDKELQMTGAENLYRATSNRVR 67
QY 121 ETVALELYSVNSNLQILKEELELSGGVDPGRHSGEAVTVPMIPGLKTKELDWSLTK 180
DB 68 ETVALELYSVNSNLQILKEELELSGGVDPGRHSGEAVTVPMIPGLKTKELDWSLTK 127
QY 181 ELISVHFGEDGASYEAEIREALRQAMRTPSRNSGLELLTAYYNQLCFDARELTTPAR 240
DB 128 ELISVHFGEDGASYEAEIREALRQAMRTPSRNSGLELLTAYYNQLCFDARELTTPAR 187


```

QY 241 SLGLFFHWYDSLTVPAQORALAFKESVLFNIGALHTQIGARODRSCTEGARRAMEAFQ 300
DB 188 SLGLFFHWYDSLTVPAQORALAFKESVLFNIGALHTQIGARODRSCTEGARRAMEAFQ 247
QY 301 RAAGAFSLLENFNSHAPSPDMSAASLCALEQLMMAQAQECVFEGLSPSPASMAPQDCLAQL 360
DB 248 RAAGAFSLLENFNSHAPSPDMSAASLCALEQLMMAQAQECVFEGLSPSPASMAPQDCLAQL 307
QY 361 RLAEQAAQVAAEYRLVHRTMAQPPVHDYVPSWTALVHVKAEBYFRSLAHYHVAMALCDGS 420
DB 308 RLAEQAAQVAAEYRLVHRTMAQPPVHDYVPSWTALVHVKAEBYFRSLAHYHVAMALCDGS 367
QY 421 PATEGELPTHQVLOPFTSSKPRGPVLPQLEERRQLGKAHLKRAILGQBEALRLHALC 480
DB 368 P-TEGELPTHQVLOPFTSSKPRGPVLPQLEERRQLGKAHLKRAILGQBEALRLHALC 426
QY 481 RVLREVDLLRAVISQTLQSLAKYAELDREDDFCEAAEAPDIQPKTHQKPEARMRLSQG 540
DB 427 RVLREVDLLRAVISQTLQSLAKYAELDREDDFCEAAEAPDIQPKTHQKPEARMRLSQG 486
QY 541 KGPDIFFHRLGLSVFSKAKRWRLVGPVHLTRGEGGFGTLRGDSVPLIAAVIPGQAAAA 600
DB 487 KGPDIFFHRLGLSVFSKAKRWRLVGPVHLTRGEGGFGTLRGDSVPLIAAVIPGQAAAA 546
QY 601 GLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAEAGASLQVWSLLPSSRLPSLGDRRPVLL 660
DB 547 GLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAEAGASLQVWSLLPSSRLPSLGDRRPVLL 606
QY 661 GPRGLLRSOREHGCKTPASTWASPRPLLNWSRKAQOGKGTGGCPQCAPVKPAPPSLKH 720
DB 607 GPRGLLRSOREHGCKTPASTWASPRPLLNWSRKAQOGKGTGGCPQCAPVKPAPPSLKH 666
QY 721 GMP 723
DB 667 GMP 669

RESULT 5
US-10-697-266-4
; Sequence 4, Application US/10697266
; Publication No. US20040067522A1
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001099-CIP-DIV2
; CURRENT APPLICATION NUMBER: US/10/697,266
; CURRENT FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: 10/274,878
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-697-266-4

```

```

Query Match 61.8%; Score 2331; DB 12; Length 634;
Best Local Similarity 67.0%; Pred. No. 2.7e-195;
Matches 473; Conservative 49; Mismatches 112; Indels 72; Gaps 6;

QY 1 MILEERPDCAGAGEESPRQLISRRKPRKTRVSLRREGRLDVCRAWRCGRVHRGSDPVR 60
DB 1 MILEERPDCAGAGEESPRQDD-----GSIR----- 26
QY 61 VHVGMNPOLHVGCDLSLTQIQCGQLQSRRAQTHQIQDKELQMTGAENLYRATSNRVR 120

```

```

DB 27 -----KGYGSFVQNPQQLQSHRARLHQIISKELMRMTGAENLYRATSNRVR 74
QY 121 ETVALELSYNSNLQQLKELEELSGVDPGRHSGEAVTVPMIPLGLKETKELDOWSTPLK 180
DB 75 ETVALELSYNSNLQQLKELEELSTVDVDPGEIGITIMPLPLGLKETKELDOWATPLK 134
QY 181 ELISVHFGEDGASYAEIRELEALRQAMRTSPRNSGLELLETAYYNQCFIDARPLTPAR 240
DB 135 ELISVHFGEDGTSFETIQELEDRLQATRTSPRDEAGLDLLAAVYSQCLFIDARFSPSR 194
QY 241 SLGLFFHWYDSLTVPAQORALAFKESVLFNIGALHTQIGARODRSCTEGARRAMEAFQ 300
DB 195 SPGLLFFHWYDSLTVPAQORALAFKESVLFNIGALHTQIGARODRSCTEGARRAMEAFQ 254
QY 301 RAAGAFSLLENFNSHAPSPDMSAASLCALEQLMMAQAQECVFEGLSPSPASMAPQDCLAQL 360
DB 255 RAAGAFSLLENFNSHAPSPDMSAASLCALEQLMMAQAQECVFEGLSPSATPDICPDQL 314
QY 361 RLAEQAAQVAAEYRLVHRTMAQPPVHDYVPSWTALVHVKAEBYFRSLAHYHVAMALCDGS 420
DB 315 QLAQEAQVAAEYGLVHRTMAQPPVHDYVPSWTALVHVKAEBYFRSLAHYHVAMALCDGS 374
QY 421 PATEGELPTHQVLOPFTSSKPRGPVLPQLEERRQLGKAHLKRAILGQBEALRLHALC 480
DB 375 PA-KGELARQSHVF-QPSTPHEPLGPTLPQHPEDRRKLAHLKRAILGQBEALRLHALC 432
QY 481 RVLREVDLLRAVISQTLQSLAKYAELDREDDFCEAAEAPDIQPKTHQKPEARMRLSQG 540
DB 433 RVLKRVLDLQVVTQALRSIAKYSQLEDDDFEATEAPDIQPKTHQTP 483
QY 541 KGPDIFFHRLGLSVFSKAKRWRLVGPVHLTRGEGGFGTLRGDSVPLIAAVIPGQAAAA 600
DB 484 -----GPLSVFSTKNRWQLVGPVHMTREGGFGTLRGDSVPLIAAVIPGQAAESA 534
QY 601 GLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAEAGASLQVWSLLPSSRLPSLGDRRPVLL 660
DB 535 GLKEGDYIVSVNGQPCRWWRHAEVVTQALRSIAKYSQLEDDDFEATEAPDIQPKTHQTP 634
QY 661 GPRGLLRSOREHGCKTPASTWASPRPLLNWSRKAQOGKGTGGCPQPC 706
DB 595 W-----NQRECGFETPMPTRTPWPILGWSRKNKQGTGSHDPDC 634

```

```

RESULT 6
US-10-274-878-4
; Sequence 4, Application US/10274878
; Publication No. US20030049792A1
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001099-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,878
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-274-878-4

```

```

Query Match 61.8%; Score 2331; DB 14; Length 634;
Best Local Similarity 67.0%; Pred. No. 2.7e-195;
Matches 473; Conservative 49; Mismatches 112; Indels 72; Gaps 6;

QY 1 MILEERPDCAGAGEESPRQLISRRKPRKTRVSLRREGRLDVCRAWRCGRVHRGSDPVR 60

```

```
Db      1 MILERPDQGGTGESSRRPQDD-----GSIR-----26
QY      61 VHVGMNPOLHAGVCDLSLTQIQCGLOQSRAAQHQIQKELQMRGTGAENLYRATSNRVR 120
Db      27 -----KGYGSFVQNPQGLQSHRARLHQIQSKELMRGTGAENLYRATSNRVR 74
QY      121 ETVALELSVNSNLQKLEELSGVDPGRHGSEAVTVPMIPLGKTKELDWSPTPLK 180
Db      75 ETVALELSVNSNLQKLEELASTSVDPQPEGEGITPMIPLGKTKELDWSPTPLK 134
QY      181 ELISVHFEGDGSYAEIRELEALQAMRTPSRNSESGLLELTAYYNQLCFDARFLTPAR 240
Db      135 ELISEHFGEDGTSFETEIOELEDLRAQTTPSRDEAGLDLLAAYISQLCFDARFSPSR 194
QY      241 SLGLFFHWYDSLTGVPAQORALAFKSGVLFNIGALHTQIGARQDCSCTEGTNHAAEAQ 300
Db      195 SPGLLFHWYDSLTGVPAQORALAFKSGVLFNIGALHTQIGARQDCSCTEGTNHAAEAQ 254
QY      301 RAAGAFSLRLRENFSHAPSPDMSAASLCALQOLMMAQAQCEVFEGLSPASMAPQDCLOL 360
Db      255 RAAGAFRLRLRENFSHAPSPDMSAASLSMLEQLMIAQAQECIFKGLLLPASATPDICPDQL 314
QY      361 RLAAQAAQVAAEYRLVHRTMAQPPVHDYVPVSWTALVHVKAHYFRSLAHYHVAMALCDGS 420
Db      315 QLAQAAQVATYEGLVHRAQAQPPVRDYLPASWTNLAHVKAHEHFCALAHYHAMALCESH 374
QY      421 PATEGELPHEQVTFOPPTSSKPRGVPVLPOELERROLGKAHLKRAILQOEALRLHALC 480
Db      375 PA-KGELARQEHVF-QPSTPHEPLGPTLPQHPEDRRKLAKAHLKRAILQOEALRLHTLC 432
QY      481 RVLRVOLLRAVISQTLORSIAKYAELDREDDPCAAEAAPDQPKTHQKPEARMPLRSOG 540
Db      433 RVLRKVOLLQVVVTQALRRSLAKYSQLEREDDFFATEAPDIQPKTHQTE-----483
QY      541 KGPDFHRLGLPSVSAKNRMLVGPVHLTRGEGGFGTLRGDSVPLIAAVIPGSOAAA 600
Db      484 -----GPSVSTKNRWQLVGPVHMTREGGFGFTLRGDSVPLIAAVIPGGOAES 534
QY      601 GLKGGDIYVSVNGQPCRWHRHAEVVTTELKAAAGASLOVSLPSSRLPSLGDREPVLL 660
Db      535 GLKGGDIYVSVNGQPCRWKHLVVTQLRSMGEEGVSLQVSLPSPPEPRGTGPRAALL 594
QY      661 GPRGLLRSOREHGCKTPASTWASPRLLNWSRKAQCKTGCGPQC 706
Db      595 W-----NORECGFETPMPTRTRPWPILGWSRKNKQKGTGSHDPDC 634

RESULT 7
US-10-239-607-71
; Sequence 71, Application US/10239607
; Publication No. US20030219761A1
; GENERAL INFORMATION:
; APPLICANT: Saatoglu, Fahri
; TITLE OF INVENTION: No. US20030219761A1el Prostate-Specific or
; TITLE OF INVENTION: Testis-Specific Nucleic Acid Molecules, Polypeptides, and
; TITLE OF INVENTION: Diagnostic and Therapeutic Methods
; FILE REFERENCE: 50218/003002
; CURRENT APPLICATION NUMBER: US/10/239,607
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09410
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,929
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 643
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-239-607-71

Query Match      61.8%; Score 2331; DB 15; Length 643;
```

```
Best Local Similarity 67.0%; Pred. No. 2.8e-195;
Matches 473; Conservative 49; Mismatches 112; Indels 72; Gaps 6;
QY      1 MILERPDQGGTGESSRRPQDD-----GSIR-----26
Db      1 MILERPDQGGTGESSRRPQDD-----GSIR-----26
QY      61 VHVGMNPOLHAGVCDLSLTQIQCGLOQSRAAQHQIQKELQMRGTGAENLYRATSNRVR 120
Db      27 -----KGYGSFVQNPQGLQSHRARLHQIQSKELMRGTGAENLYRATSNRVR 74
QY      121 ETVALELSVNSNLQKLEELSGVDPGRHGSEAVTVPMIPLGKTKELDWSPTPLK 180
Db      75 ETVALELSVNSNLQKLEELASTSVDPQPEGEGITPMIPLGKTKELDWSPTPLK 134
QY      181 ELISVHFEGDGSYAEIRELEALQAMRTPSRNSESGLLELTAYYNQLCFDARFLTPAR 240
Db      135 ELISEHFGEDGTSFETEIOELEDLRAQTTPSRDEAGLDLLAAYISQLCFDARFSPSR 194
QY      241 SLGLFFHWYDSLTGVPAQORALAFKSGVLFNIGALHTQIGARQDCSCTEGTNHAAEAQ 300
Db      195 SPGLLFHWYDSLTGVPAQORALAFKSGVLFNIGALHTQIGARQDCSCTEGTNHAAEAQ 254
QY      301 RAAGAFSLRLRENFSHAPSPDMSAASLCALQOLMMAQAQCEVFEGLSPASMAPQDCLOL 360
Db      255 RAAGAFRLRLRENFSHAPSPDMSAASLSMLEQLMIAQAQECIFKGLLLPASATPDICPDQL 314
QY      361 RLAAQAAQVAAEYRLVHRTMAQPPVHDYVPVSWTALVHVKAHYFRSLAHYHVAMALCDGS 420
Db      315 QLAQAAQVATYEGLVHRAQAQPPVRDYLPASWTNLAHVKAHEHFCALAHYHAMALCESH 374
QY      421 PATEGELPHEQVTFOPPTSSKPRGVPVLPOELERROLGKAHLKRAILQOEALRLHALC 480
Db      375 PA-KGELARQEHVF-QPSTPHEPLGPTLPQHPEDRRKLAKAHLKRAILQOEALRLHTLC 432
QY      481 RVLRVOLLRAVISQTLORSIAKYAELDREDDPCAAEAAPDQPKTHQKPEARMPLRSOG 540
Db      433 RVLRKVOLLQVVVTQALRRSLAKYSQLEREDDFFATEAPDIQPKTHQTE-----483
QY      541 KGPDFHRLGLPSVSAKNRMLVGPVHLTRGEGGFGTLRGDSVPLIAAVIPGSOAAA 600
Db      484 -----GPSVSTKNRWQLVGPVHMTREGGFGFTLRGDSVPLIAAVIPGGOAES 534
QY      601 GLKGGDIYVSVNGQPCRWHRHAEVVTTELKAAAGASLOVSLPSSRLPSLGDREPVLL 660
Db      535 GLKGGDIYVSVNGQPCRWKHLVVTQLRSMGEEGVSLQVSLPSPPEPRGTGPRAALL 594
QY      661 GPRGLLRSOREHGCKTPASTWASPRLLNWSRKAQCKTGCGPQC 706
Db      595 W-----NORECGFETPMPTRTRPWPILGWSRKNKQKGTGSHDPDC 634

RESULT 8
US-10-408-765A-2465
; Sequence 2465, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2465
; LENGTH: 311
```


; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/158,003
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 87
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-896-87

Query Match 33.3%; Score 1278; DB 14; Length 257;
Best Local Similarity 99.2%; Pred. No. 1.9e-103;
Matches 245; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 459 GKAKLRAILGQEAALRLHALCRVLREVDLLRAVISQTLQRSIAKVAELDDDFCEAAE 518
Db 8 GKAKLRAILGQEAALRLHALCRVLREVDLLRAVISQTLQRSIAKVAELDDDFCEAAE 67
QY 519 APDIQPKTHQKPEARMRLSOGKGPDIHFHRLGPLSVESAKNRRLVGPVHLTRGEGFGL 578
Db 68 APDIQPKTHQKPEARMRLSOGKGPDIHFHRLGPLSVESAKNRRLVGPVHLTRGEGFGL 127
QY 579 TLRGDSFVLIAAIPGSQAAAAAGLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAAGEAGASL 638
Db 128 TLRGDSFVLIAAIPGSQAAAAAGLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAAGEAGASL 187
QY 639 QVSVLLPSSRLPSLGRDRPVLLGPRGLLRQRHCKTTPASTWASPRPLNWSRKAQOQK 698
Db 188 QVSVLLPSSRLPSLGRDRPVLLGPRGLLRQRHCKTTPASTWASPRPLNWSRKAQOQK 247
QY 699 TGGCPOP 705
Db 248 TGGCPOP 254

RESULT 12
US-10-108-260A-4840
; Sequence 4840, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; FILE OF INVENTION: No. US20040005560A1 full length cdna
; TITLE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4840
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4840

Query Match 33.3%; Score 1256; DB 15; Length 686;
Best Local Similarity 45.9%; Pred. No. 7.3e-101;
Matches 266; Conservative 107; Mismatches 196; Indels 10; Gaps 5;

QY 74 GDSLTQIQGQLQSRRAHQIQDKELQVKTGAENLYRATSNRVRVETVALELSYVNSN 133
Db 24 GCNPLAQTKRSKLQNGRAALNQILKAMRMRTGAENLLKVATNSKVRQVRLSELSFVNSD 83
QY 134 LQMLELELSGVDPGRHSGEAVTVPMPLGLKETKELDWSGTPKELISVHFGEAGAS 193
Db 84 LQMLELELSGVDPGRHSGEAVTVPMPLGLKETKELDWSGTPKELISVHFGEAGAS 143
QY 194 YEAEIIELEALQAMTTPRNEGLELLTAYYNQLCFLDARFLTPARSLGPEHWDLSLT 253
Db 144 YEDEIADMLDQACHTPDRDEAGVELLMTYFICLGFVESRFPPTQRMGLLFTWYDSUT 203
QY 254 GYPAQORALAFKSGSVLFNIGALHTQIGARQDRSCTEGARRAMEAFQRAAGAFSLRENF 313
Db 204 GYVPSQONLLEKASVLFNTGALYTGITGTCDFQTOAGLESAIDAFQRAAGVNLVLTDF 263

QY 314 SHAPSPDMSAASICALQELMMAQAQECVFEGLSPSPMAPQDCIAQLRLAQEAQAQVAAEY 373
Db 264 THTPSYDMSPAMLVSLVKVMAQAQESVFEKISLPG--IRNEFFMLVKVQAQAQVAAEY 321
QY 374 RLVRHTMAQPPVHDYVSVSWTALVHVKAQEFYFRSLAHYHVMALCDG--SPATEGELPTE 431
Db 322 QQLHAAMSQAQPVKENIPYSWASLACVKAHHAALAHYFTAILLIDHQVKPFTDID--HQ 378
QY 432 QVFLQPTSSKPRG--PVLIP-QELEERROLGKAKLRAILGQEAALRLHALCRVLREVDL 488
Db 379 EKCLSQLYDEMPGLTFLATLKNQDQRRQLGKSHLRRAVHAHEESVREASCLKLSIEV 438
QY 489 LRAVISQTLQRSIAKVAELDDDFCEAAEAPDIQPKTHQKPEARMRLSOGKGPDIHFR 548
Db 439 LQKVLCAAQERSRLTYAQHQQEEDDLNLDAPSVAKTEQEVDIILPQFSKLTVDFFQK 498
QY 549 LGLSVFSKAKNRRLVGPVHLTRGEGFGLTRGDSFVLIAAIPGSQAAAAAGLKEGDYI 608
Db 499 LGLSVFSKAKNRRLVGPVHLTRGEGFGLTRGDSFVLIAAIPGSQAAAAAGLKEGDYI 558
QY 609 VSVNGQPCRWWRHAEVVTTELKAAAGEAGASLQVSVLLPSS 647
Db 559 VSIQLVDCNKLTLSEVWMLKLSFGDEIEKVVSLDST 597

RESULT 13
US-10-092-900A-352
; Sequence 352, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Liu, Yi
; APPLICANT: Anderson, David W.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Catterton, Elina
; APPLICANT: Leite, Mario W.
; APPLICANT: Zhong, Haihong
; APPLICANT: Alsbrook, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-290C
; CURRENT APPLICATION NUMBER: US/10/092,900A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: USSN 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/283,675
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: USSN 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: USSN 60/274,281


```
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: Aeonica Sequence Listing Engine
; SEQ ID NO 3
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-895-040A-3

Query Match      33.1%; Score 1249; DB 9; Length 686;
Best Local Similarity 45.8%; Pred. No. 3e-100;
Matches 265; Conservative 107; Mismatches 197; Indels 10; Gaps 5;

QY      74  GDSLTIQIQGQLQSRRAQHQQIDKELQWRTGAENLYRATSNRVRETVALELSYVNSN 133
Db      24  GCNPLAQTGRSLQNGEAALNQILKAVNRNIGAEKLLKVATNSKVRQVRLELSFVNSD 83

QY      134 LQLLKEELSELSCGVDGRHGSBAVTPMIPGLKXETKELDWSTPLKELISVHFGEDEGAS 193
Db      84  LQMLKEELEGLNTSVGVYQNTBEAFTIPLIPLGLKETKQVDPAVLKDFILEHYSDEGYL 143

QY      194 YEAREIELEALROAMETPSRNESGLELLTAYYNQLCFDLDFLTTPARSIGLFFHWYDSLT 253
Db      144 YEDEIADLMDLRQACRTPSRDEAGVELLMTYFIQLGFVESRFPPTQMGLLFTWYDSLT 203

QY      254 GVPAQORALAFEGKSVLFNIGALHTOIGARQDRSCTEGAERAMEAFORAAAGAFSLLRENF 313
Db      204 GVPVSQQNLLLEKASVLENTGALYTQIGTCRQTOAGLESADAFQRAAGVLYNLYKDTF 263

QY      314 SHAPSPDMGAASLCALQLEQLMAQAQCEVFEGLSPSPASMAPQDCIAQLRLAQEAQAAVEY 373
Db      264 THTPSYDMSFAMLSVLVKMMLAQAEQSVFEKISLPG--IRNBEFMLVKVAQEAQKGEVY 321

QY      374 RLVRHRTMAQPPVEDYVYVSWTALVHVKAQYFRSLAHVHVAMALCDG--SPATEGELPTHE 431
Db      322 QQLHAAMSQAQPVKENIPYSNASLACVKAHHAALAHYFTAILLIDHQVKPGTDLD--HQ 378

QY      432 QVFLQPPTSSKPRG--PVLP-QELERRQIQGAHLKRAILQGEALRLHALCRVLREVDL 488
Db      379 EKCLSQLYDHMPPEGLTPLATKNDQQRQLGKSHLRAMAHHEESVREASLCKLRSIEV 438

QY      489 LRAVISQTLQRSIAKVAELDREDDFCEAAAPADIPKTHQKPEARMPLRSQGGKPDIFHR 548
Db      439 LQKVLCAQAQERSRLTYAQHEEDDLNLDAPSWAKTEQEVDIILPQFSKLVITDFFQK 498

QY      549 LGPLSVFSANRWLVGPVHLTREGGFGLTIRGDSPLVIAAVIPGSQAAAAGLKEGDYI 608
Db      499 LGPLSVFSANKRWTPPRSRIRFTAEAGDLGFTLRGNAPVQVHFILDYPCSAVAGAREGDYI 558

QY      609 VSVNGQPCRWRHAEVVTTELKAAEGASLQVVSLLPSS 647
Db      559 VSIQLVDCRWLTISEVMKLLKSFGEDEIEMKVVSLLDST 597
```

Search completed: July 13, 2004, 12:22:09
Job time : 429 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2004, 08:26:41 ; Search time 25 Seconds
(without alignments)

1505.869 Million cell updates/sec

Title: US-10-697-266-2

Perfect score: 3774

Sequence: 1 MILEERPDCGAGEBPRLO.....QPCAPVKPAPPSLKHPGW 723

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3419	90.6	695	1 RHNI_HUMAN	Q8tcx5 mus sapien
2	2331	61.8	643	1 RHNI_MOUSE	Q61085 mus musculu
3	284.5	7.5	861	1 PQS8_CABEL	P34552 caenorhabdi
4	274.5	7.3	868	1 PD61_HUMAN	Q8wmu4 homo sapien
5	266.5	7.1	869	1 PD61_MOUSE	Q9wu78 mus musculu
6	261.5	6.9	867	1 PD61_XENLA	Q9w6c5 xenopus lae
7	202	5.4	844	1 BR01_YEAST	P48582 saccharomyc
8	156.5	4.1	942	1 PKL1_HUMAN	Q16512 homo sapien
9	146.5	3.9	946	1 PKL1_RAT	Q63433 rattus norv
10	146	3.9	1017	1 EM11_MOUSE	Q99x41 mus musculu
11	145.5	3.9	1522	1 ARHB_HUMAN	Q15085 homo sapien
12	138	3.7	910	1 US1C_MOUSE	Q9es64 mus musculu
13	136.5	3.6	1527	1 ARHB_RAT	Q9es67 rattus norv
14	135.5	3.6	1016	1 EM11_HUMAN	Q9y6c2 homo sapien
15	132	3.5	1544	1 ARHC_HUMAN	Q9nzn5 homo sapien
16	131.5	3.5	838	1 KFC2_HUMAN	Q96ac6 homo sapien
17	131.5	3.5	2564	1 SPQ2_HUMAN	Q9h254 homo sapien
18	130	3.4	1387	1 RGSC_RAT	Q08774 rattus norv
19	129	3.4	1205	1 SMC_MYCTU	Q10970 mycobacteri
20	128.5	3.4	1455	1 AIPI_HUMAN	Q86ul8 homo sapien
21	128.5	3.4	3674	1 SPCK_HUMAN	Q9nrc6 homo sapien
22	128	3.4	1543	1 ARHC_MOUSE	Q8r4h2 mus musculu
23	127.5	3.4	4684	1 PLE1_HUMAN	Q15149 homo sapien
24	126	3.3	4687	1 PLE1_RAT	P30427 rattus norv
25	124	3.3	701	1 YAB5_SCHPO	Q09807 schizosacch
26	124	3.3	1275	1 AIPI_MOUSE	Q9wyq1 mus musculu
27	124	3.3	4473	1 PLE1_CRIGR	Q9j155 cricetus
28	123.5	3.3	1816	1 LMA4_HUMAN	Q16363 homo sapien
29	121.5	3.2	1277	1 AIPI_RAT	Q88382 rattus norv
30	121	3.2	817	1 MUT5_THECA	Q9zix6 thermus cal
31	121	3.2	1447	1 RGSC_HUMAN	Q14924 homo sapien
32	119.5	3.2	818	1 MUT5_THETH	Q56239 thermus the
33	119	3.2	1920	1 PCT2_MOUSE	P48725 mus musculu

34	119	3.2	5327	1 MACF_MOUSE	Q9qxx0 mus musculu
35	118	3.1	330	1 RIL_CHICK	Q9pw72 gallus gall
36	118	3.1	651	1 PRIA_MYCLE	Q9ccg3 mycobacteri
37	118	3.1	1445	1 PTPG_HUMAN	P23470 homo sapien
38	117	3.1	668	1 FAU_DROME	Q9vqx3 drosophila
39	116.5	3.1	531	1 GL1A_ORYSA	F47997 oryza sativ
40	116.5	3.1	1739	1 CHD2_HUMAN	O14647 homo sapien
41	116	3.1	946	1 RHG4_HUMAN	P98171 homo sapien
42	116	3.1	1238	1 SBCC_RHOCA	O68032 rhodobacter
43	116	3.1	2442	1 CEP2_HUMAN	Q9bw73 homo sapien
44	115.5	3.1	655	1 PRIA_MYCTU	P71670 mycobacteri
45	115.5	3.1	933	1 ZO3_HUMAN	O95049 homo sapien

ALIGNMENTS

RESULT 1
RHNI_HUMAN
ID RHNI_HUMAN STANDARD; PRT; 695 AA.
AC Q8TCX5;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Rhophilin 1.
GN RHPN1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Burbelo P.D.;
RT "Structure and function of rhopilin homologs."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Binds specifically to GTP-Rho. May serve as a target for
CC Rho. It has no enzymatic activity. It may interact with some
CC cytoskeletal component upon Rho binding or relay a Rho signal to
CC other molecules (By similarity).
CC -!- SIMILARITY: Contains 1 BR01 domain.
CC -!- SIMILARITY: Contains 1 HR1 domain.
CC -!- SIMILARITY: Contains 1 PDZ/DHR domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AY082588; AAL89809.1; -;
CC Genew; HGNC:19973; RHPN1.
CC InterPro; IPR004328; BR01.
CC InterPro; IPR001478; PDZ.
CC InterPro; IPR000861; REM_repeat.
CC Pfam; PF03097; BR01; 1.
CC Pfam; PF02185; HR1; 1.
CC Pfam; PF00595; PDZ; 1.
CC SMART; SM00074; HR1; 1.
CC SMART; SM00228; PDZ; 1.
CC PROSITE; PS50106; PDZ; 1.
CC FT DOMAIN 35 107 HR1.
CC FT DOMAIN 108 260 BR01.
CC FT DOMAIN 538 617 PDZ.
CC SQ SEQUENCE 695 AA; 76292 MW; 89E4DF6806761452 CRC64;

Query Match 90.6%; Score 3419; DB 1; Length 695;
Best Local Similarity 89.4%; Pred. No. 1.1e-218;
Matches 669; Conservative 0; Mismatches 1; Indels 78; Gaps 2;

QY 1 MILEERPDCGAGEBPRLOISRRKPRKTRVSSLRGRREGULRDVCAWRGCRVHRGEDPVR 60
|||||

```
Db 1 MILLERPDGAGAGESPRLQ----- 20
Qy 61 VHVGMNQLHAGVCDSTLQCCGLOSRRAQIHQIDKELQMTGAENLYRATSNRVR 120
Db 21 -----CGDSTLQCCGLOSRRAQIHQIDKELQMTGAENLYRATSNRVR 67
Qy 121 ETVALELSVNSNLQLLKEELEELSGGVDPRHGSEAVTVPMIPLGLKETKELDWSTPLK 180
Db 68 ETVALELSVNSNLQLLKEELEELSGGVDPRHGSEAVTVPMIPLGLKETKELDWSTPLK 127
Qy 181 ELISVHFGEDGASVEAEIRELEALRQAMRTSPSRNESGLELLTAYYNQCLFDARFLTPAR 240
Db 128 ELISVHFGEDGASVEAEIRELEALRQAMRTSPSRNESGLELLTAYYNQCLFDARFLTPAR 187
Qy 241 SLGLFFHWYDSLTGVPQAORALAFKGVLFNIGALHTQIGARQDRSCTEGARRAMEAFQ 300
Db 188 SLGLFFHWYDSLTGVPQAORALAFKGVLFNIGALHTQIGARQDRSCTEGARRAMEAFQ 247
Qy 301 RAAGAFSLLRENFHSHAPSMDMSAASLCALQLEMMQAQCEVPEGLSPSPASMAPQDCLAOL 360
Db 248 RAAGAFSLLRENFHSHAPSMDMSAASLCALQLEMMQAQCEVPEGLSPSPASMAPQDCLAOL 307
Qy 361 RLQAEAAQVAAYEIRLVHRTMAQPPVHDVYVPSVSWTALVHVKAAYFRSLAHYHVAMALCDGS 420
Db 308 RLQAEAAQVAAYEIRLVHRTMAQPPVHDVYVPSVSWTALVHVKAAYFRSLAHYHVAMALCDGS 367
Qy 421 -----PATEGELPTHEQVFLQPTSSKPRGPVLPQLEER 455
Db 368 RECPLHFMVLPRPPRAGSQPLCPAAYEGELPTHEQVFLQPTSSKPRGPVLPQLEER 427
Qy 456 RQLGKAHLKRALIGQEEALRHLCRLVREVDLLRAVLSQTLQSLAKYAELEDREDDFCE 515
Db 428 RQLGKAHLKRALIGQEEALRHLCRLVREVDLLRAVLSQTLQSLAKYAELEDREDDFCE 487
Qy 516 ABAEDIPQKTHQKPEARNPRLSQKGPDI FRLGLPVSFSNAKRWLVGPVHLTRGEGG 575
Db 488 ABAEDIPQKTHQKPEARNPRLSQKGPDI FRLGLPVSFSNAKRWLVGPVHLTRGEGG 547
Qy 576 FGLTLRGDSPVLIAAVIPGSAAGLKEGDYIVSVNGQPCRWRRHAEVVTCLKAAGBAG 635
Db 548 FGLTLRGDSPVLIAAVIPGSAAGLKEGDYIVSVNGQPCRWRRHAEVVTCLKAAGBAG 607
Qy 636 ASLVQVSLPSSRLPSLGRDRPVLLGPRGLLSQREHGCKTPASTWASPRPLLNWSRKAQ 695
Db 608 ASLVQVSLPSSRLPSLGRDRPVLLGPRGLLSQREHGCKTPASTWASPRPLLNWSRKAQ 667
Qy 696 QGKTGGCPQCAPVKPAPPSSLKHPGW 723
Db 668 QGKTGGCPQCAPVKPAPPSSLKHPGW 695
```

RESULT 2

```
RHNI_MOUSE
ID RHNI_MOUSE STANDARD; PRG; 643 AA.
AC Q61085;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Rhophilin 1 (GTP-Rho binding protein 1).
GN RHPN1 OR GRBP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Embryo;
RX MEDLINE=96165390; PubMed=8571126;
RA Watanabe G., Saito Y., Madaule P., Ishizaki T., Fujisawa K., Morii N.,
RA Mukai H., Ono Y., Kakizuka A., Narumiya S.;
RT "protein kinase N (PKN) and PKN-related protein rhophilin as targets
RT of small GTPase Rho.";
RL Science 271:645-648(1996).
```

```
CC -!- FUNCTION: Binds specifically to GTP-Rho. May serve as a target for
CC Rho. It has no enzymatic activity. It may interact with some
CC cytoskeletal component upon Rho binding or relay a Rho signal to
CC other molecules.
CC -!- TISSUE SPECIFICITY: Highly expressed in testis.
CC -!- SIMILARITY: Contains 1 BRO1 domain.
CC -!- SIMILARITY: Contains 1 PDZ/DHR domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcements
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U43194; AAC52388.1; -.
CC HSSP: P29476; IQAV.
CC MGD: MGI:1098783; Rbpnl.
CC GO: GO:0005083; F:small GTPase regulatory/interacting protein. .; IPI.
CC InterPro: IPR004328; BRO1.
CC InterPro: IPR001478; PDZ.
CC InterPro: IPR000861; REM_repeat.
CC Pfam: PF03097; BRO1; 1.
CC Pfam: PF02185; HRI; 1.
CC Pfam: PF00595; PDZ; 1.
CC SMART: SM00074; HRI; 1.
CC SMART: SM00228; PDZ; 1.
CC PROSITE: PS50106; PDZ; 1.
CC DOMAIN 42 114 HRI.
CC FT DOMAIN 115 267 BRO1.
CC FT DOMAIN 500 577 PDZ.
CC SEQUENCE 643 AA; 71288 MW; 5CD24E1417D8C6D8 CRC64;
Qy Query Match 61.8%; Score 2331; DB 1; Length 643;
Qy Best Local Similarity 67.0%; Pred. No. 9.7e-147; Indels 72; Gaps 6;
Qy Matches 473; Conservative 49; Mismatches 112;
Db 1 MILLERPDGAGAGESPRLQISRRKPRKTRVSSLSRREGRLDVCAMRGCRVHRGEDPVR 60
Db 1 MLLEBPDGQGTGESSRPQDD-----GSIR----- 26
Qy 61 VHVGMNQLHAGVCDSTLQCCGLOSRRAQIHQIDKELQMTGAENLYRATSNRVR 120
Db 27 -----KGYGSFVQNPQGLQSHRARLHQIISKELAMRTCAENLYRATSNRVR 74
Qy 121 ETVALELSVNSNLQLLKEELEELSGGVDPRHGSEAVTVPMIPLGLKETKELDWSTPLK 180
Db 75 ETVALELSVNSNLQLLKEELAEELSTSDVDQPEGEGITIPMLPLGLKETKELDWATPLK 134
Qy 181 ELISVHFGEDGASVEAEIRELEALRQAMRTSPSRNESGLELLTAYYNQCLFDARFLTPAR 240
Db 135 ELISVHFGEDGTSFETIQLLEDLAQATRTPSRDEAGLDLLAAYISQLCFLDARFSPSR 194
Qy 241 SLGLFFHWYDSLTGVPQAORALAFKGVLFNIGALHTQIGARQDRSCTEGARRAMEAFQ 300
Db 195 SPGLLFHWYDSLTGVPQAORALAFKGVLFNIGALHTQIGARQDRSCTEGTHAAAEAFQ 254
Qy 301 RAAGAFSLLRENFHSHAPSMDMSAASLCALQLEMMQAQCEVPEGLSPSPASMAPQDCLAOL 360
Db 255 RAAGAFRLLENFHSHAPSMDMSAASLSMLEQLMIAQAQECIFKGLLLPASATPDICPDOL 314
Qy 361 RLQAEAAQVAAYEIRLVHRTMAQPPVHDVYVPSVSWTALVHVKAAYFRSLAHYHVAMALCDGS 420
Db 315 QLQAEAAQVATEYGLVHRMAQPPVRDYLPAASWTNLAVHVKAESHFCALAHYHMAALCESH 374
Qy 421 PATEGELPTHEQVFLQPTSSKPRGPVLPQLEERROLGKHLKRALIGQEEALRHLC 480
Db 375 PA-KGELAQEHPV-OPSTPHEPLGPTLPQHPEDRRKLAKHLKRALIGQEEALRHLC 432
Qy 481 RVLRVDLLRAVLSQTLQSLAKYAELEDREDDFCEAAEAPDIQPKTHQKPEARNPRLSQG 540
Db 481 RVLRVDLLRAVLSQTLQSLAKYAELEDREDDFCEAAEAPDIQPKTHQKPEARNPRLSQG 540
```

Db 433 RVLKRVLDLQVVVTVQALRRSLAKYSQLEERDDFFATEAPDIQPKTHOTPE----- 483
QY 541 KGPDI FHLRGPLSVFSAXNRWRLVGPVHLTRCEGGFGLTIRGDSPLVIAAIVPGSQAATA 600
Db 484 -----GPLSVFSTKNRWQLVGPVHMTREGGFGFTIRGDSPLVIAAIVPGSQAESA 534
QY 601 GLKEGDYIVSVNGQPCRWHRHAEVVTTELKAAEAGASLQVSLPSRLPSIGDRRPVLL 660
Db 535 GLKEGDYIVSVNGQPCRWKHLVTVTLRSMGEEGVSLQVSLPSPEPRGTGPRRAALL 594
QY 661 GPRGLLRSOREHGCKTPASTWASPRLLNWSKAQOGKGGCPQPC 706
Db 595 W-----NORECGFETPMPTRTTRPWPILGWSRXKQOGKGGSHPDPC 634

RESULT 3
ID P058 CABEL STANDARD; PRT; 861 AA.
AC P34552; Q94159;
DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Protein pgn-58 (Protein YNK1).
GN PQN-58 OR R10E12.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=98087404; PubMed=9427532;
RA Che S., Weil M.M., Etkin L.D., Epstein H.F., Kuang J.;
RT "Molecular cloning of a splice variant of Caenorhabditis elegans YNK1,
a putative element in signal transduction."
RL Biochim. Biophys. Acta 1354:231-240 (1997).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=79063198;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fraser A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lighthouse J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkneen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans."
RL Nature 368:32-38 (1994).
[3]
RN REVISIONS.
RA Durbin R.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Probably involved in signal transduction.
CC -!- SIMILARITY: TO MAMMALIAN GRBP; SOME, TO YEAST BROL.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U73679; AAC67305.1; -
CC DR EMBL; Z29561; CAAB2667.2; -
CC DR WormPep; R10E12.1; CE25075.
CC InterPro; IPR004328; BROL.

DR Pfam; PF03097; BROL1; 1.
FT DOMAIN 756 861 GLN-RICH.
SQ SEQUENCE 861 AA; 96245 MW; F2F600623FF74965 CRC64;
Query Match 7.5%; Score 284.5; DB 1; Length 861;
Best Local Similarity 22.8%; Pred. No. 3.7e-11;
Matches 132; Conservative 98; Mismatches 209; Indels 139; Gaps 22;
QY 167 LKETELDKSTPLKELIS--VHFGDGSAYEAE-TRELEAIR-QAMRT-PRNESGLELL 221
Db 11 LKSTNEVDLVKPLTSYIDNVNTSDNRRSDVAEAVQELNKLKSKACCQCLDQHQSALDVL 70
QY 222 TAYYNQLCFDLARFLTPARSGLGFHWHYDSL-TGVPAAQORAL-----AFEXKGSVLFNIG 274
Db 71 TRYDOLVAIENKIITISATQNPVFNKDKDFPKGSLFSSRASLSLSDGSFERAAVLFNIG 130
QY 275 ALHTQIGARQDRCTEGARRAMEAFQRAAGAFSLRE----NFSHAPSDDMSAASICAL 330
Db 131 SLMSQIGAAQQFHTDDEIKVSAKLFOQSAGVFARLRDVVLGMVQOEFTPLDMPDTILAALS 190
QY 331 QLMMAQAQECVFEGLSPASMAPQDCLAQRLAQEAQAQVAAEYRLVHRTWAQPPVHDYVP 390
Db 191 ALMTAQAEAIY-----IKGHEKMKATSMVKISAQVAEFYSEAKQMKSKDIVRGLMD 243
QY 391 VSMTALVHVKAERYSLAHVHVAMALCDGSPATEGELPTHEQVFLQPTTSKPRGPVLPQ 450
Db 244 KWSALVSGKULAIQALAQYHQS-EVCG-----
QY 451 ELEERRQLGKAHLKRAILGQEEALRLHALCRVLRVDLLRAVISQ-----TLQSLAK- 503
Db 271 ---EARQIGE-----QLSRLAESLKLFDTAQKYLPRDITGWIDYPSVSKA 313
QY 504 YALREDDEPCEAAEAPDQPKTHQKPEARMPRLSQGKP-----DIEHRLGPLSVESA 557
Db 314 HAAAKKDNDFIYHEKVSDFR-TLPTLPKAVIAKPTMQTPTMTPTSFDFMFAVLPVQVHNA 372
QY 558 KRW-----RLVGPVHLTRGEGGGLTLRGDSPVLIAAIVPGSQAATAAGLKE 604
Db 373 MQSYDARKAELVNMTVRMEATQLMNG-----VLASLNLPAAL----- 411
QY 605 GDYIVSVNGQPCRWHRHAEVVTTELKAA--GEAGASLQVSLPSRLPSIGDRRPVLLGP 662
Db 412 -DDVTSFETLP-----ESLKLKSAKLQNGSGSSIMRLF--SELPTLYQRNEDILTE 460
QY 663 RGLLSOREHGCKTPASTWASPRLLNWSKAQOGKGTG 700
Db 461 TSRILNEEKESDDTMRKQLGT-----KWTMRSSSQLTG 493

RESULT 4

PD61 HUMAN
ID PD61 HUMAN STANDARD; PRT; 868 AA.
AC Q8WUM4; Q9BX86; Q9NUN0; Q9P2H2; Q9UKL5;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Programmed cell death 6-interacting protein (PDCD6-interacting
DE protein) (ALG-2 interacting protein 1) (Hp95).
GN PDCD6IP OR AIP1 OR KIAA1375.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN SEQUENCE FROM N.A.
RX MEDLINE=21533686; PubMed=11683497;
RA Wu Y., Pan S., Che S., He G., Nelman-Gonzalez M., Weil M.M., Kuang J.;
RT "Overexpression of Hp95 induces G1 phase arrest in confluent HeLa
cells."
RL Differentiation 67:139-153 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Li H., Shioda T., Isselbacher K.J.;

RT "Molecular cloning of human ALG-2 interacting protein 1 (AIP1).";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
 RA Koundinya M., Raphael J., Moreira D., Kelley T., Labaer J., Lin Y.,
 RA Phelan M., Farmer A.,
 RT "Cloning of human full-length CDSs in BD Creator(TM) system donor
 vector.";
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lymph;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg K.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.-I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Brakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP SEQUENCE OF 209-868 FROM N.A.
 RC TISSUE=Placenta;
 RA Isoqai T., Ota T., Hayaishi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Watanabe S., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saio K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
 RA Ninomiya K., Iwayanagi T.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 323-868 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20181126; PubMed=10718198;
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XVI.
 The complete sequences of 150 new cDNA clones from brain which code
 for large proteins in vitro.";
 RL DNA Res. 7:65-73(2000).
 CC -!- FUNCTION: May play a role in the regulation of both apoptosis and
 CC cell proliferation.
 CC -!- SUBUNIT: Interacts with PDCD6/ALG-2 and SH3BP1 (By similarity).
 CC -!- SIMILARITY: Contains 1 BRO1 domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF349951; AA020398.1; -
 CC EMBL; AF151793; AA082220.1; -
 CC EMBL; BT007367; AAP36031.1; -
 CC EMBL; BC020066; AA020066.1; -
 CC EMBL; AK002122; BAA32092.1; ALT_INIT.

DR EMBL; AB037796; BAA92613.1; -
 DR Genew; HGNC:8766; PDCD6IP.
 DR MIM; 608074; -
 DR InterPro; IPR004328; BRO1.
 DR Pfam; PF03097; BRO1; 1.
 KW Apoptosis.
 FT DOMAIN 1 167 BRO1.
 FT PRO-RICH.
 FT 717 860
 FT CONFLICT 309 309 A -> T (IN REF. 2).
 FT CONFLICT 378 378 V -> I (IN REF. 5).
 FT CONFLICT 550 550 N -> S (IN REF. 1 AND 6).
 FT CONFLICT 580 580 M -> T (IN REF. 5).
 FT CONFLICT 730 730 S -> L (IN REF. 2).
 SQ SEQUENCE 868 AA; 96023 MW; 573588D1F612EC93 CRC64;
 Query Match 7.3%; Score 274.5; DB 1; Length 868;
 Best Local Similarity 27.2%; Pred. No. 1.7e-10;
 Matches 116; Conservative 57; Mismatches 152; Indels 101; Gaps 17;
 QY 163 IPLGLKETKELDWSTPLKELISVHF---GEGDGSAYEAEIRELEALRQAM--RTPSRNESC 217
 DB 5 ISVQLKKTSEVDLAKELVFIQOTYPSGGEGQAQYCRAAEEELSKLRAAVRGRLDKHEGA 64
 QY 218 LELLTAYNQLCLDARFLTPARSLGLFFHWYD-----SLTG--VPAQORALAFKGSYL 270
 DB 65 LETLLRYYPQICSIIBPKFPFSENQICLTFTWKADPKGSLFGSVKALASLYEKS CVL 124
 QY 271 ENIGALHTOIGARODRSCTEGHARRAMEAFQRAAGFSLREN----FSHAPSMDMSAASL 326
 DB 125 FNCALASQIAAQNLDNDEGLKIAKHVQFASGAFHLIKETVLSALSREPTVDISPTDV 184
 QY 327 CALEOLMMAQACQECVFEGLSPSPASMAPQDCLAQLRAQEAQAQAAEYRLVHRTMAQPPVH 386
 DB 185 GTLSLIMLAQAQEVFF-----LKATRDMDKALIAKLANQ--AADY--FGDAFKOCQYK 234
 QY 387 DYVPVSVTALVHVKAERYFSLAHYHVALCDGSPATEGELTHEQVFLQPTSSKPRGS 446
 DB 235 DTLP-----KEVPVLAAKHCIM-----QANAETHQSI-----262
 QY 447 VLPQLEBERQLGKHLKRAILGQEARLRLHLCRVLEVDLLRAV-----TS 494
 DB 263 -----LAKQKK---FGESIALRLOHA-----AELIKTVASGYDEVVNVKDFS 301
 QY 495 QTLQSLAKYAELEDDDDFCBAEAAPDIQ-----PKTHQKPEARMPLRSQCKGPDIFHR 548
 DB 302 DKINEAL---AAAKNDNFVHYDRVPLDKLDLPICKATLVKSTPVNVPISQ-KFTDLPEK 357
 QY 549 LGPLSV 554
 DB 358 MVPVSV 363
 RESULT 5
 PDC1_MOUSE
 ID PDC1_MOUSE STANDARD; PRT; 869 AA.
 AC Q9WU78; O88695; O89014; O8BSL8; O8R0H5; Q99LR3; Q9CZNR;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Programmed cell death 6 interacting protein (ALG-2 interacting protein
 DE X) (ALG-2 interacting protein 1) (E2F1-inducible protein) (Big2).
 GN PDCD6IP OR ALIX OR AIP1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND INTERACTION WITH PDCD6.
 RC TISSUE=Brain;
 RX MEDLINE=99218669; PubMed=10200558;
 RA Missorten M., Nichols A., Rieger K., Sadoul R.;
 RT "Alix, a novel mouse protein undergoing calcium-dependent interaction
 with the apoptotic-linked-gene (ALG-2) protein.";

Cell Death Differ. 6:124-129(1999).
 [2]
 RN SEQUENCE FROM N.A. (ISOFORM 1).
 RP MEDLINE=99098896; PubMed=9880530;
 RX Vito P., Pellegrini L., Guet C., D'Adamo L.;
 RT "Cloning of Aip1, a novel protein that associates with the apoptosis-
 linked gene ALG-2 in a Ca2+-dependent reaction.";
 RL J. Biol. Chem. 274:1533-1540(1999).
 [3]
 RN SEQUENCE FROM N.A. (ISOFORM 3).
 RP STRAIN=C57BL/6J; TISSUE=Forelimb;
 RX MEDLINE=22354683; PubMed=12466851;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaiko I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Knadt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalia E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konegaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 [4]
 RN SEQUENCE FROM N.A. (ISOFORM 1).
 RP TISSUE=Breast;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshlyuk S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchan J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smalius D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [5]
 RN SEQUENCE OF 671-869 FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Forelimb;
 RX MEDLINE=20027348; PubMed=10556317;
 RA Wang A.J., Pierce A., Judson-Kremer K., Gaddis S., Aldaz C.M.,
 RA Johnson D.G., MacLeod M.C.;
 RT "Rapid analysis of gene expression (RAGE) facilitates universal
 expression profiling.";

Nucleic Acids Res. 27:4609-4618(1999).
 -!- FUNCTION: May play a role in the regulation of both apoptosis and
 cell proliferation.
 -!- SUBUNIT: Interacts with PDCD6/ALG-2 and SH3BP1 (By similarity).
 CC The interaction with PDCD6 is dependent on calcium.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=Q9WU78-1; Sequence=Displayed;
 CC Name=2; Synonyms=Alx-SP, Short;
 CC IsoId=Q9WU78-2; Sequence=VSP 007502;
 CC Note=Does not interact with ALG-2;
 CC Name=3;
 CC IsoId=Q9WU78-3; Sequence=VSP 007501;
 CC -!- TISSUE SPECIFICITY: Ubiquitously expressed.
 CC -!- SIMILARITY: Contains 1 BRO1 domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).

 EMBL; AJ005073; CAA06329.1; -
 EMBL; AJ005074; CAA06330.1; -
 EMBL; AF119955; AAD26813.1; -
 EMBL; AK031256; BAC27323.1; -
 EMBL; BC002261; AAH02261.1; -
 EMBL; BC026823; AAH26823.1; -
 EMBL; AF176514; AAD53115.1; -
 MGD; MGI:133753; Pdc6ip.
 GO; GO:0005829; C:cytosol; IDA.
 InterPro; IPR004328; BRO1.
 Pfam; PF03097; BRO1; 1.
 Apoptosis; Alternative splicing.
 DOMAIN 1 167
 FT DOMAIN 717 861
 FT VARSPPLIC 239
 FT
 FT VARSPPLIC 159 805
 FT
 FT CONFLICT 329 333
 FT CONFLICT 530 530
 FT CONFLICT 547 548
 FT CONFLICT 595 595
 FT CONFLICT 625 625
 FT CONFLICT 640 641
 FT CONFLICT 821 821
 FT CONFLICT 853 853
 FT CONFLICT 853 853
 SQ SEQUENCE 869 AA; 96010 MW; 9AA84B592FDCEASE CRC64;
 Query Match 7.1%; Score 266.5; DB 1; Length 869;
 Best Local Similarity 26.1%; Pred. No. 5.9e-10;
 Matches 110; Conservative 60; Mismatches 151; Indels 101; Gaps 15;
 QY 167 LKRYKELDWSTPLKELISVHF---GEGASYEAEIRRELEALRQAM--RTPSRNKGLELL 221
 Db 9 LKKTSEVDLAKELVKFIQOTVPSGEEQAYCRAAEELSKLRSSALGPLDKHGALETL 68
 QY 222 TAYYNQLCFDLARFLTPARSIGLFFHWYD-----SLTG--VPAQORALAFKSGVLFNIG 274
 Db 69 LRYVDQICSIIEPKFPFSENQICLTFTWKDAFDKGLFGSVKVLASLGYEKSVCVFNCA 128
 QY 275 ALHTQIGARQDRSCTEGARAMEAFORAGAFSLREN----FSHAPSPDMSASLCALE 330
 Db 129 ALASQIAAEQNLDNDEGLKTAQKQYQFASGAFLEHKOTVLSALSREPTVDISPDVTGLS 188
 QY 331 OLWMAQAQCEVFEGLSPDASMAPQDCIAQLRLAQEAQVAEYRLVHRTMAQPPVHDYVP 390
 Db 189 LIMLAQAQEVTF-----LKAIRDKMDALIAKLANO-AADY---FGDAFKQCQYKOTLP 238

QY 391 VSWTALVHVKAEPFRSLAHYHVMALCDGSPATEGELPTEHQVFLQPTSSKRGVLPQ 450
Db 239 KEVPTLAAQCIQMAAAYHQSI-----262
QY 451 ELBERROLGKAHUKRAILGQEAHLHLCVRLREVLLRAV-----ISOTLQ 498
Db 263 -----LAKQOKX---FGBEIARLOHA-----AELIKNVASRYDEYVNVKDFSKIN 305
QY 499 RSLAKYAELOREDDFCEAAEAPDIQ-----PKTHQKPEARMPLRSOGKGPDIHRLGPL 552
Db 306 RAL---TAARKNDNFIYHDRVVDLKDLPDGKATLVKPTVNVPSQ-KFTDLFEKVPV 361
QY 553 SV 554
Db 362 SV 363

RESULT 6

ID PD6I XENLA STANDARD; PRT; 867 AA.
AC Q9W6C5;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Programmed cell death 6 interacting protein (Signal transduction protein Xp95).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxID=8355;
[1]
FN SEQUENCE FROM N.A. (ISOFORM 1), AND PHOSPHORYLATION.
RP MEDLINE=99150334; PubMed=10026166;
RA Che S., El-Hodiri H.M., Wu C.-F., Nelman-Gonzalez M., Weil M.M.,
RA Etkin L.D., Clark R.B., Kuang J.;
RT "Identification and cloning of Xp95, a putative signal transduction protein in Xenopus oocytes."
RL J. Biol. Chem. 274:5522-5531(1999).
[2]
FN SEQUENCE FROM N.A. (ISOFORM 2).
RP TISSUE=Embryo;
RA Klein S., Strausberg R.L.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May be a signal transduction protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9W6C5-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9W6C5-2; Sequence=VSP 007503;
CC -!- PTM: Phosphorylated on tyrosine residues.
CC -!- SIMILARITY: Contains 1 BR01 domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF115497; AAD20341.1; -
DR EMBL; BC043849; AAH43849.1; ALT_INIT.
DR InterPro: IPR004328; BR01.
DR Pfam; PF03097; BR01; 1.
KW Phosphorylation; Alternative splicing.
FT DOMAIN 1 166
FT PRO-RICH.
FT DOMAIN 716 866
FT VARSPLIC 238 238 K -> KYFYFQ (in isoform 2).
FT /FTIQ=VSP 007503.
SQ SEQUENCE 867 AA; 96198 MW; 6BE7173E68263B3 CRC64;

Query Match 6.9%; Score 261.5; DB 1; Length 867;
Best Local Similarity 24.5%; Pred. No. 1.3e-09;
Matches 106; Conservative 72; Mismatches 144; Indels 111; Gaps 16;
QY 158 VTVPMTPLGKKEYKELDWSTPLKELISVHF--GEDGASYEAEIRELEALQAM--RTPSR 213
Db 5 ISVP-----LKKYSEVDLVKPLSKYIHNTYPSGSDQTEYCRAVDELNKLRSASVGRPLDK 59
QY 214 NESGELLTAYYNQCLFCDARFLTTPARSLGLFPHWYDSLTL-----GVPAQQRALAFEK 266
Db 60 HETSLETFWERYDOLCSVEPKFPFTESQLCTITWKDAFDKGSIFGGSVKLALPSLGVK 119
QY 267 GSVLFNIGALHTQIGARQDRSCTEGARRAMEAFORAGAFSLIRE----NFSHAPSMDMS 322
Db 120 TCVLFNIGALASQIASQNLNDNEALKASKFYQLASGAFSHIKDTVTLSSLNRPDPTVDIS 179
QY 323 AASLCALFQLMMAQAEQCVFEGLSPPASMAPODCLQRLAQEAQAQVAAEYRLVHRTMAQ 382
Db 180 PDIVGTLSLIMLAQAEVFF-----LKATRDQKDAVIKLANQA-----220
QY 383 PPVHDYVPSVSWTALVHVKAEPFRSLAHYHVMALCDGSPATEGELPTEHQVFLQPTSSK 442
Db 221 ----DY-----YGDFAKQCQ-----YKDTLSK 238
QY 443 PRGVLPQELERROLGKAHLK--RAILGQEAHLHLCVRLREVLLRAV-----492
Db 239 EVFPI---LAAKHCIMQAHAHVQSLAKQOKKFGEEIGRLQHASLIVKTVSSRYDEYV 294
QY 493 ----ISOTLQSLAKYAELOREDDFCEAAEAPDI-----QPKTHQKPEARMPLRSOGK 541
Db 295 NVKDLADKINRAL---TAARKNDNFIYHDRVVDLKDLPDGKASLVKSTPVNVF-LSQ-K 349
QY 542 GPDIFHRLGPLSV 554
Db 350 YTDLFEKVPVAV 362

RESULT 7

BR01 YEAST
ID BR01 YEAST STANDARD; PRT; 844 AA.
AC P48582; 002823;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE BR01 protein.
GN BR01 OR YPL084W OR LPF2.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96220442; PubMed=8649366;
RA Nickas M.E., Yaffe M.P.;
RT "BR01, a novel gene that interacts with components of the
RT Pkc1p-mitogen-activated protein kinase pathway in Saccharomyces
RT cerevisiae."
RL Mol. Cell. Biol. 16:2585-2593(1996).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=97313271; PubMed=9169875;
RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansong W.,
RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
RA Dietrich F.S., Dalius H., DiPaolo T., Dubois E., Duesterhoeft A.,
RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.W.,
RA Hunnicke-Smith S., Hyman R., Johnston M., Kalman S., Klein K.,
RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
RA Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,
RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,

NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
MEDLINE=95080426; PubMed=7988719;
Palmer R.H., Ridden J., Parker P.J.;
"Identification of multiple, novel, protein kinase C-related gene products."; FEBS Lett. 356:5-8(1994).
[2]
SEQUENCE FROM N.A.
TISSUE=Fetal brain;
MEDLINE=95154310; PubMed=7851406;
Palmer R.H., Ridden J., Parker P.J.;
"Cloning and expression patterns of two members of a novel protein kinase-C-related kinase family."; Eur. J. Biochem. 227:344-351(1995).
[3]
SEQUENCE FROM N.A., AND MUTAGENESIS OF LYS-644.
TISSUE=Hippocampus;
MEDLINE=94183274; PubMed=8135837;
Mukai H., Ono Y.;
"A novel protein kinase with leucine zipper-like sequences: its catalytic domain is highly homologous to that of protein kinase C. Biochem. Biophys. Res. Commun. 199:897-904(1994).
-!- FUNCTION: CAN PHOSPHORYLATE RIBOSOMAL PROTEIN S6. MEDIATES GTP- RHO DEPENDENT INTRACELLULAR SIGNALLING (BY SIMILARITY).
-!- ENZYME REGULATION: Activated by lipids, particularly cardiolipin and to a lesser extent by other acidic phospholipids (By similarity).
-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
-!- TISSUE SPECIFICITY: Found ubiquitously. Expressed in heart, brain, placenta, lung, skeletal muscle, kidney and pancreas.
-!- PM: Autophosphorylated; preferably in serine.
-!- PPM: Activated by limited proteolysis with trypsin (By similarity).
-!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL, out of the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/a> or send an email to license@isb-sib.ch).

EMBL; U33053; AAC50209.1; -;
DR EMBL; S75546; AAB33345.1; -;
DR EMBL; D26181; BAA05169.1; -;
DR PIR; JC2129; JC2129.
DR PDB; 1CX2; 23-MAY-00.
DR Genew; HGNC:9405; PRKCL1.
DR MIM; 601032; -;
DR GO; GO:0004672; P:protein kinase activity; TAS.
DR GO; GO:0007257; P:activation of JUNK; TAS.
DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CaLB.
DR InterPro; IPR000961; Pkinase C.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR000861; REM repeat.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF02185; HRI; 3.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00433; pkinase C; 1.
DR ProDom; PD000001; prot_kinase; 1.
DR SMART; SM00239; C2; 1.
DR SMART; SM00074; HRI; 3.
DR SMART; SM00133; S_TK_X; 1.
DR SMART; SM00220; S_TK; 1.
DR DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.

DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 KW Transferase; ATP-binding; Serine/threonine-protein kinase;
 FT Phosphorylation; Polymorphism; 3D-structure.
 FT DOMAIN 615 874 PROTEIN KINASE.
 FT NP BIND 621 629 ATP (BY SIMILARITY).
 FT BINDING 644 644 ATP (BY SIMILARITY).
 FT ACT SITE 740 740 BY SIMILARITY.
 FT VARIANT 901 901 V->I (in dSNP:10846).
 FT /FTID=VAR_014937.
 FT MUTAGEN 644 K-R: SUBSTANTIAL REDUCTION OF
 FT AUTOPHOSPHORYLATION.
 FT CONFLICT 191 191 D->G (IN REF. 3).
 SQ SEQUENCE 942 AA; 103989 MW; A89B40DCABF560E3 CRC64;

Query Match 4.1%; Score 156.5; DB 1; Length 942;
 Best Local Similarity 21.4%; Pred. No. 0.012;
 Matches 159; Conservative 79; Mismatches 257; Indels 247; Gaps 37;

QY 69 QLHVGCDSLTIQCGQLQGRRAIQHOIDKELQMRGAENLYRATSN-NRVRETVALEL 127
 Db 18 QLGLAGADLAAPGVQQQLERERLRREIRKELKKEGAENLRATTDLGSLGPEVLL 77
 QY 128 SYVNSNIQLLKEELLEISGGV---DP-GRHGEAVTVPMPLGLKTKELDMWSTPLKELI 183
 Db 78 RGSRRLLDLHLQQLQELHAAHVLPDPAATHDG-----PQSP----- 113
 QY 184 SVHFGEGASYEA-EIRELEALRQAMRTPSNESGLELLTAYYNQLCFDLARFILTARS 242
 Db 114 ----GAGGPTCSATNLSRVAGLEKQLALELVKQGAENMIQYTSNGSTKDRKLLTAQOM 169
 QY 243 GLFPHWYDSLTVGP-----AQRALAFE-----KGSVLFNIGALHTQIGA----- 282
 Db 170 -----LQDSKTKIDIIIMQLRRAQLQADLENOAAPDQTQSP--DLGAVELRIEELRHHF 222
 QY 283 RQDRSCTEGARAMEAFQRAAGAFSLIRENFHAPSMDMSAASICALQLMMAQAQBCVF 342
 Db 223 RVEHAVAEGAKNVLRL-----LSAAKAPDRKAVSEAAQEKLATESNOKLGLLR 268
 QY 343 EGLSPFASMAPDQCLAQRLAQEAQVAEYELVHRTWAQPPVHDY----VPVSWTALVH 398
 Db 269 EALERLIGELPAD-HPKGRLRLRELAASAAFTSLRGLAPPATHYSTLCKPAPLTGTL 327
 QY 399 VKAEYFRSLAH---YHVAMALCDGSPATEGELPTHEQVFLQPP----- 438
 Db 328 VRVVGCRDLPETIPWNPFSM--GGPGI-----PDSRPFLSRPARGLYSRGSLSGRSL 381
 QY 439 -----TSSKPRGP-----VLPQLEERRQLGKAHLKRAILGQ 470
 Db 382 KAAEANTSEVSTVLKLDNTVVGQTSWKPCGPNAMDQSFTELEARE-----LELAVFWR 436
 QY 471 EEARLHALC--RVLREVDLRAVISQTLQSLAKYAEILDREDDFCEAAEAPDIQPKTHQ 528
 Db 437 DO-----RGLCALFKLJEDFL-----DNERHEVDLMEPEQCLVAEV-----TFR 477
 QY 529 KPE-ARMPRL-----SQCKGPDIFHRLGFLSVFSKAKRWRLVGPVHLTR----- 571
 Db 478 NPVIERIPLRRQKKIFSKQQGA---FQRAQNMIDVA--TW-----VRLRLRLIPNAT 527
 QY 572 GEGGFLTLRGDSPVLIAAVTPGSAAGLKEGDYIVSVNGQPCRWWRHAEVVTLEKAA 631
 Db 528 GTGTF-----SP-----GASPGSARTTG-----DISVE-----KINLG 556
 QY 632 GEAGASLOWVSLPLSSRLPSLGDRLPVLGPRGLLRSQREHGCKTPASTWASPRLLNWS 691
 Db 557 TDSDSPQKSRDPSPSSSLSS--PI-----QESTAPELFSEQTETFGPAL--- 601
 QY 692 RKAQQKGTGGCPCQPCAPVKPAP 713
 Db 602 -----CSPLRKS 609

PKL1 RAT
 ID PKL1 RAT STANDARD; PRT; 946 AA.
 AC Q63433;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Protein kinase C-like 1 (EC 2.7.11.-) (Protein-kinase C-related kinase
 DE 1) (Protein kinase C-like PKN) (Serine-threonine protein kinase N)
 DE (Protease-activated kinase 1) (PAK-1).
 GN PRKCL1 OR PRK1 OR PKN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=94183274; PubMed=8135837;
 RA Ono Y., Mukai H.;
 RT "A novel protein kinase with leucine zipper-like sequences: its
 RT catalytic domain is highly homologous to that of protein kinase C.";
 RL Biochem. Biophys. Res. Commun. 199:897-904(1994).
 RN [2]
 RP CHARACTERIZATION.
 RC TISSUE=Liver;
 RX MEDLINE=94327556; PubMed=8051089;
 RA Morrice N.A., Gabrielli B., Kemp B.E., Wettenhall R.E.;
 RT "A cardiolipin-activated protein kinase from rat liver structurally
 RT distinct from the protein kinases C.";
 RL J. Biol. Chem. 269:20040-20046(1994).
 CC -!- FUNCTION: CAN PHOSPHORYLATE RIBOSOMAL PROTEIN S6. MEDIATES GTPASE
 CC RHO DEPENDENT INTRACELLULAR SIGNALING.
 CC -!- ENZYME REGULATION: Activated by lipids, particularly cardiolipin
 CC and to a lesser extent by other acidic phospholipids.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- PM: Autophosphorylated; preferably in serine.
 CC -!- PM: Activated by limited proteolysis with trypsin.
 CC -!- SIMILARITY: Belongs to the Ser/thr family of protein kinases. PKC
 CC subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D26180; BAA05168.1; -;
 CC HSSP; P05132; 1ATP.
 CC InterPro; IPR000008; C2.
 CC InterPro; IPR008973; C2 CalB.
 CC InterPro; IPR000961; Pkinase_C.
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR000861; RM_repeat.
 CC InterPro; IPR008271; Ser_thr_pkin_AS.
 CC InterPro; IPR002290; Ser_thr_pkinase.
 CC Pfam; PF02185; HR1; 3.
 CC Pfam; PF00069; pkinase; 1.
 CC Pfam; PF00433; pkinase_C; 1.
 CC ProDom; P000001; Prot_kinase; 1.
 CC SMART; SM00239; C2; 1.
 CC SMART; SM00074; HR1; 3.
 CC SMART; SM00133; S_TK_X; 1.
 CC SMART; SM00220; S_TK_X; 1.
 CC PROSITE; PS00108; PROTEIN KINASE ST; 1.
 CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 CC Transpherase; ATP-binding; Serine/threonine-protein kinase;
 KW Phosphorylation.
 FT DOMAIN 615 874 PROTEIN KINASE.
 FT NP BIND 625 633 ATP (BY SIMILARITY).
 FT BINDING 648 648 ATP (BY SIMILARITY).

CC nucleation of the triple helix and then a further quaternary
 CC assembly to higher order polymers via intermolecular disulfide
 CC bonds. Interacts with EMILIN2 (By similarity).
 CC -!- SUBCELLULAR LOCATION: Extracellular matrix. Found mainly at the
 CC interface between amorphous elastin and microfibrils.
 CC -!- DEVELOPMENTAL STAGE: Detectable in morula and blastocyst. First
 CC expressed in ectoplacental cone in embryos of 6.5 days and in
 CC extraembryonic visceral endoderm at 7.5 days. Expressed also in
 CC the allantois. Expression in the ectoplacental cone-derived
 CC secondary trophoblast giant cells and spongiotrophoblast is strong
 CC up to 11.5 days and then declines. In the embryo, high levels are
 CC initially expressed in blood vessels, perineural mesenchyme and
 CC somites at 8.5 days. Later on, intense expression is identified in
 CC the mesenchymal component of organs anlage (ie lung and liver) and
 CC different mesenchymal condensations (ie limb bud and branchial
 CC arches). At late gestation expression is widely distributed in
 CC interstitial connective tissue and smooth muscle cell-rich
 CC tissues.
 CC -!- SIMILARITY: Contains 1 C1Q domain.
 CC -!- SIMILARITY: Contains 1 collagenous domain.
 CC -!- SIMILARITY: Contains 1 EMI domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AK029337; AAC26403.1; -;
 CC EMBL; BC005481; AAH05481.1; -;
 CC MGD; MGI:1926189; Emilin1.
 CC InterPro; IPR001073; C1q.
 CC InterPro; IPR008160; Collagen.
 CC InterPro; IPR008993; TNF_like.
 CC Pfam; PF00386; C1q; 1.
 CC Pfam; PF01391; Collagen; 1.
 CC Cell adhesion; Extracellular matrix; Signal; Coiled coil; Collagen;
 KW Glycoprotein. 1 23
 FT SIGNAL 24 1017 POTENTIAL.
 FT CHAIN 57 131 EMILIN 1.
 FT DOMAIN 171 211 COILED COIL (POTENTIAL).
 FT DOMAIN 237 266 COILED COIL (POTENTIAL).
 FT DOMAIN 269 275 POLY-GLY.
 FT DOMAIN 310 374 COILED COIL (POTENTIAL).
 FT DOMAIN 519 573 COILED COIL (POTENTIAL).
 FT DOMAIN 676 697 COILED COIL (POTENTIAL).
 FT DOMAIN 789 809 COILED COIL (POTENTIAL).
 FT DOMAIN 815 865 COLLAGEN.
 FT DOMAIN 873 1008 C1Q.
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 416 416 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 562 562 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 659 659 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 767 767 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 795 795 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SEQUENCE 1017 AA; 107584 MW; 2EF903204DB9C8F CRC64;
 SQ
 Query Match 3.9%; Score 146; DB 1; Length 1017;
 Best Local Similarity 21.0%; Pred. No. 0.069;
 Matches 183; Conservative 87; Mismatches 277; Indels 324; Gaps 45;
 47 WRGCRVHGED-----PVRVHVGPMNPQLHANGCDSLTQICCG 85
 120 WRCCQYGGDCCGEGPASVLGPAPSTPLPRPRPRVRLNSGSSAGSLGLGGGPGVESEK 179
 86 LQSRRAQIHQIDKELQMTG-----AENLYRATS-----VRETVALE 126
 180 VQQLERQV-KSLTKELQGLRGVLQGMNGLAEDVQRADVTFVNGRQOPADAAARPGVHET 238

QY 127 LSYVNSNLQLL-----KEELEEL-----SGVDPRHSGSEAVTVPMIPGLGKETKELDW 175
 Db 239 LSEIQOQLLDNRVSTHDQELGHLNNHHNGG--PGGGGRASGPVP--VPSGPSEELLRL 295
 QY 176 STPLKELISVHF-CEDGASYEAE-----IRELEALQAM----- 208
 Db 296 ERQLQSCSVCLTGLDGFROQQQDRERLRTLEKLMSSMEERQQQLVCPAMARRPPQSCC 355
 QY 209 -----RTPSRNESGLELL----- 235
 Db 356 PPELGRVSELELRDLDVVTGSLTVLSGRGSELGGAAGGGHPGYTSLASRLSLEDRF 415
 QY 236 -----LTPARSL-----GLFFHWYDLSITGPAQORALAFKGVSLFNGALHTOIGARQ 284
 Db 416 NSTLGPSEBEQKNWPGGRLCHW---LPAAPGRLEKLE---GLLANVS---RELGRM 465
 QY 285 DRSCTEGARAMEAFQRAAGAFSLLENFNSHAPSMDMAAS--LCALQLQMM----- 334
 Db 466 D-----LLE--EQVAGAVRTCGQICSGAPGEQDSRVNEILLSALERRVLDSEGRQL 514
 QY 335 -----AOAQECVFEGSLPPSPASMAPQDCLAQLAQLAQA-AQVAAEYRL-VHRTMA 381
 Db 515 VSGGLHEAAEAGEAQAQVLEGL-----QGLLSRLRERMDAQOETAAILRLNLTA 566
 QY 382 QPPVHDYVVPVWTALVHVKAIEFRSLAHYVAMALCDGSPATEGEL-----PTHEQ 432
 Db 567 QLS-----OLEGLLQARGD-----EGCGACGGVQBELGLRLDRGVKSCPCP--- 606
 QY 433 VFLQPP-----TSSKPRGVPVLQPELEERRQLGKAHLKRAILGQEEALRLHALCRVLR 484
 Db 607 --LLPPRGAGPGVGGSRGPL-----DGFSVFGGSSGSAQLAQGELS 649
 QY 485 EVDLLRAVISQTLQSLAKYAELEDDDFCEAAEAPDIQPKTHQKPEARMPLRSOGKGP 544
 Db 650 EVILTFSSINDSLH----- 675
 QY 545 IPRHLGPL--SVFSAKNRWRLVGPVHLTRGEGGF-----GLTLRGDSPVL----- 587
 Db 676 LAD-LGATKDSIISINLQOEATEHVTESEERFRGLEEGQAQGCPSLEGRLGLEGV 734
 QY 588 ---IAAVIPGSOAAAAGLKEGDYIVSVNGQPCRWW---RHAERVTELKAA-----GEA 634
 Db 735 CERLOTVAGGLQ---GLREG-----LSHVAVGLAAVAVRESNSTSLTQAALKLLGGQA 785
 QY 635 GASLQVSVLLPSRLPSLGDRLPVL-----LQPRGLLSQRSHGCKTPTASTWASPLPLN 689
 Db 786 GLGRRIGAL--NNSLLLEDLRLQSLKDTGPGSKAGPPGPPGLQGGSPGAPGPP--- 841
 QY 690 WSRKAQOGKTGCPQP-----CAPVKEAP 713
 Db 842 -GKDGQQAIGP-PGPGQGAEGAPAPVP 870
 RESULT 11
 ID ARHB HUMAN STANDARD; PRT; 1522 AA.
 AC O15085;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Rho guanine nucleotide exchange factor 11 (PDZ-RhoGEF).
 GN ARHGEF11 OR KIAA0380.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97349984; PubMed=9205841;
 RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. VII."

RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro.";
 RL DNA Res. 4:141-150(1997).
 RN [2]
 RP INTERACTION WITH RHOA.
 RX MEDLINE=99456836; PubMed=10526156;
 RA Ruemennapp U., Blomquist A., Schwoerer G., Schablowski H., Psoma A.,
 RA Jakobs K.H.;
 RT "Rho-specific binding and guanine nucleotide exchange catalysis by
 RT KIAA0380, a dbl family member.";
 RL FEBS Lett. 459:313-318(1999).
 RN [3]
 RP INTERACTION WITH GNA12 AND GNA13, AND TISSUE SPECIFICITY.
 RC TISSUE=Brain;
 RX MEDLINE=99150378; PubMed=10026210;
 RA Fukuhara S., Murga C., Zohar M., Igishi T., Gutkind J.S.;
 RT "A novel PDZ domain containing guanine nucleotide exchange factor
 RT links heterotrimeric G proteins to Rho.";
 RL J. Biol. Chem. 274:5868-5879(1999).
 RN [4]
 RP INTERACTION WITH PLXNB1 AND PLXNB2.
 RX MEDLINE=2225992; PubMed=12372594;
 RA Driessens M.H., Olivo C., Nagata K., Inagaki M., Collard J.G.;
 RT "B plexins activate Rho through PDZ-RhoGEF.";
 RL FEBS Lett. 529:168-172(2002).
 RN [5]
 RP INTERACTION WITH PLXNB1 AND PLXNB2.
 RX MEDLINE=22301907; PubMed=12183458;
 RA Ferrer V., Vazquez-Prado J., Gutkind J.S.;
 RT "Plexin B regulates Rho through the guanine nucleotide exchange
 RT factors leukemia-associated Rho GEF (LARG) and PDZ-RhoGEF.";
 RL J. Biol. Chem. 277:43115-43120(2002).
 RN [6]
 RP SUBCELLULAR LOCATION.
 RX MEDLINE=20449035; PubMed=10900204;
 RA Togashi H., Nagata K., Takagishi M., Saitoh N., Inagaki M.;
 RT "Functions of a rho-specific guanine nucleotide exchange factor in
 RT neurite retraction. Possible role of a proline-rich motif of KIAA0380
 RT in localization.";
 RL J. Biol. Chem. 275:29570-29578(2000).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 281-490.
 RX MEDLINE=21364065; PubMed=11470431;
 RA Longenecker K.L., Lewis M.E., Chikumi H., Gukkind J.S.,
 RA Derewenda Z.S.;
 RT "Structure of the RGS-like domain from PDZ-RhoGEF: linking
 RT heterotrimeric g protein-coupled signaling to Rho GTPases.";
 RL Structure 9:559-569(2001).
 CC -!- FUNCTION: May play a role in the regulation of RhoA GTPase by
 CC guanine nucleotide-binding alpha-12 (GNA12) and alpha-13 (GNA13).
 CC Acts as guanine nucleotide exchange factor (GEF) for RhoA GTPase
 CC and may act as GTPase activating protein (GAP) for GNA12 and
 CC GNA13.
 CC -!- SUBUNIT: Interacts with GNA12 and GNA13 through the RGS domain.
 CC Interacts with RHOA, PLXNB1 AND PLXNB2. Interacts with SLCL16 (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic. Translocated to the membrane
 CC upon stimulation.
 CC -!- TISSUE SPECIFICITY: Ubiquitously expressed.
 CC -!- DOMAIN: The poly-Pro region is essential for plasma membrane
 CC localization upon stimulation.
 CC -!- SIMILARITY: Contains 1 DBL-homology (DH) domain.
 CC -!- SIMILARITY: Contains 1 PDZ/DHR domain.
 CC -!- SIMILARITY: Contains 1 PH domain.
 CC -!- SIMILARITY: Contains 1 RGS (RGS-like) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AB002378; BAA20834.1; -.
 DR PDB; 1HTJ; 11-JUL-01.
 DR Genew; HGNC:14580; ARHGEF11.
 DR TIM; 605708; -.
 DR InterPro; IPR001331; GDS_CDC24.
 DR InterPro; IPR001478; PDZ.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR000342; Regl Gproteins.
 DR InterPro; IPR000219; RhoGEF.
 DR Pfam; PF00595; PDZ; 1.
 DR Pfam; PF00621; RhoGEF; 1.
 DR SMART; SM00228; PDZ; 1.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00315; RGS; 1.
 DR SMART; SM00325; RhoGEF; 1.
 DR PROSITE; PS00741; DH 1; FALSE_NEG.
 DR PROSITE; PS001010; DH 2; 1.
 DR PROSITE; PS0106; PDZ; 1.
 DR PROSITE; PS00003; PH DOMAIN; 1.
 DR PROSITE; PS0132; RGS; 1.
 KW Guanine-nucleotide releasing factor; GTPase activation; Coiled coil;
 KW 3D-structure.
 FT DOMAIN 47 126 PDZ.
 FT DOMAIN 306 486 RGS.
 FT DOMAIN 444 470 COILED COIL (POTENTIAL).
 FT DOMAIN 734 923 DH.
 FT DOMAIN 965 1079 PH.
 FT DOMAIN 1094 1099 POLY-PRO.
 SQ SEQUENCE 1522 AA; 167703 MW; CA16L25B9F8A4AA CRC64;
 Query Match 3.9%; Score 145.5; DB 1; Length 1522;
 Best Local Similarity 31.9%; Pred.No. 0.13;
 Matches 51; Conservative 16; Mismatches 56; Indels 37; Gaps 6;
 QY 567 VHLTRGEGGGLTLCGDSVLTAAVTPGSAANGKEGDIYVNVGQPCRWRRHAEVVT 626
 Db 48 VTIQDHGFVTSVGRVILVQSVRPGGAAMKAGVKEGDIKVKVGTWVNTSSHLEVK 107
 QY 627 ELKAGEAGASLQVWSLLPSSRLPSLGRRPVLLGRLRSQREHGCKTPASTWASRP 686
 Db 108 LIK-----SGAYV-ALTLLGSS-----PSSMGISGLQDDPSAGAPRITSVIFSP 153
 QY 687 LNWSRKAQGGTGGGCPQCAPVK-----PAP---PSSLKH 719
 Db 154 -----PPPLPPPPQRTGPKPLQDPPEVQKH 177
 RESULT 12
 USIC_MOUSE
 ID USIC_MOUSE STANDARD; PRT; 910 AA.
 AC Q9ES64; Q91XD1; Q9CVG7; Q9ES65;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Harmonin (Usher syndrome 1C protein homolog) (PDZ domain-containing
 DE protein).
 GN USH1C.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3), TISSUE SPECIFICITY, AND
 RP ALTERNATIVE SPLICING.
 RC TISSUE=Inner ear;
 RX MEDLINE=20428180; PubMed=10973247;
 RA Verpy E., Leibovici M., Zwaenepoel I., Liu X.-Z., Gal A., Salem N.,
 RA Mansour A., Blanchard S., Kobayashi I., Keats B.J.B., Slim R.,
 RA Petit C.;
 RT "A defect in harmonin, a PDZ domain-containing protein expressed in
 RT the inner ear sensory hair cells, underlies Usher syndrome type 1C.";
 RL Nat. Genet. 26:51-55(2000).


```

RN
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Verpy E., Leibovici M., Zwaenepoel I., Blanchard S., Petit C.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN
RP [3]
RC TISSUE=Colon;
RX MEDLINE=22386257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Schenck C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schaeetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
RN
RP [4]
RC SEQUENCE OF 740-852 FROM N.A. (ISOFORM 1).
RX STRAIN=C57BL/6J; TISSUE=Small intestine;
RC MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasaki Y., Kedzierski R.M., King B.L.,
RA Konggaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Nunata K., Okido T., Pavan W.J., Perlea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sadelain A., Schneider C., Sempie C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wellis C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imorani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RN
RP [5] SUBUNIT: Associates with USH1G/SANS (By similarity).
CC -!- ALTERNATIVE PRODUCTS;
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Additional isoforms exist;
CC Name=3; Synonyms=b3;
CC IsoId=Q9ES64-1; Sequence=Displayed;
CC Name=2; Synonyms=b2;
CC IsoId=Q9ES64-2; Sequence=VSP_050532, VSP_050534;
CC Name=1; Synonyms=a1;
CC IsoId=Q9ES64-3; Sequence=VSP_050530, VSP_050531, VSP_050533;
CC -!- TISSUE SPECIFICITY: Isoform 1 is expressed in the eye, cochlea,
CC vestibule, heart, kidney, small intestine and testis; it is barely
```


Db 766 LALEGGVDSFVGKVVSAVYEGGAERHGGVKGDEIMANG-----KIVTDYTLA 816
QY 632 GEAGASLQ-----VSLPLS-----SRPLSGDRRPVL-----659
Db 817 -EAAALQKAWNOGDWIDLVAVCPPEYDDELSSLPSSAABSPQLAKQLEAYEPVCR 875
QY 660 -----LGPRLRLRSQREHCKTPASTWASPRP 686
Db 876 HGFFLQLEPTLLKSKERNQTDPSWRPASSAP 908

RESULT 13

ARHB_RAT ID ARHB_RAT STANDARD; PRT; 1527 AA.
AC Q9567;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Rho guanine nucleotide exchange factor 11 (RhoGEF glutamate transport modulator GTRAP48).
GN ARHGEF11.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, AND INTERACTION WITH RHOA; GNA13 AND SLC1A6.
RX MEDLINE=21133160; PubMed=11242047;
RA Jackson M., Song W., Liu M.-Y., Jin L., Dykes-Hoberg M., Lin C.-L.G., Bowers W.J., Federoff H.J., Sternweis P.C., Rothstein J.D.;
RT "Modulation of the neuronal glutamate transporter EAAT4 by two interacting proteins";
RL Nature 410:89-93(2001).
CC -!- FUNCTION: May play a role in the regulation of Rho GTPase by guanine nucleotide-binding alpha-12 (GNA12) and alpha-13 (GNA13). Acts as guanine nucleotide exchange factor (GEF) for RhoA GTPase and may act as GTPase activating protein (GAP) for GNA12 and GNA13 (By similarity).
CC -!- SUBUNIT: Interacts with RHOA, GNA13 and SLC1A6. Interacts with GNA12, PLXNB1 and PLXNB2 (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic. Translocated to the membrane upon stimulation (By similarity).
CC -!- SIMILARITY: Contains 1 DBL-homology (DH) domain.
CC -!- SIMILARITY: Contains 1 PDZ/DHR domain.
CC -!- SIMILARITY: Contains 1 PH domain.
CC -!- SIMILARITY: Contains 1 RGS (RGS-like) domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF225961; AAG28597.1; --
HSP; Q12923; 3PDZ.
DR InterPro; IPR001331; GDS_CDC24.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000342; RegI_Gprotein.
DR Pfam; PF00595; PDZ_1.
DR Pfam; PF0621; RhoGEF_1.
DR SMART; SM00228; PDZ; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00315; RGS; 1.
DR SMART; SM00325; RhoGEF; 1.
DR PROSITE; PS00741; DH_1; FALSE_NEG.
DR PROSITE; PS50010; DH_2; 1.
DR PROSITE; PS50106; PDZ; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.

DR PROSITE; PS50132; RGS; 1.
KW Guanine-nucleotide releasing factor; GTPase activation; Coiled coil.
FT DOMAIN 64 143 PDZ.
FT DOMAIN 323 503 RGS1.
FT DOMAIN 461 487 COILED COIL (POTENTIAL).
FT DOMAIN 742 931 DH.
FT DOMAIN 973 1087 PH.
SQ SEQUENCE 1527 AA; 168533 MW; ABAEA20F541A3A9A CRC64;
Query Match 3.6%; Score 136.5; DB 1; Length 1527;
Best Local Similarity 29.2%; Pred. No. 0.51; Matches 45; Conservative 18; Mismatches 56; Indels 35; Gaps 4;
QY 567 VHLTRGEGGFLTRGDSFVLIAAVPGSOAAAGLKEGDYIVSVNGQPCRMWRHAEVVT 626
Db 65 VLIQDQHGFQFTVSGDRIVLVQSVRPGGAAMKAGVKEGDRIIKVGNGTWTNSSHLEVVK 124
QY 627 ELKAAEAGASLQVSVLLPSSRLPSLGDRPVLVLPGRLLRSQREHGCKTPTASTWASPRP 686
Db 125 LIKSGAYA-----ALTLLGSS-----PPSVGVSGLQNPVSAGV-----LRVNP 163
QY 687 LLNWSRKAQGGTGGCPQPCAPVKPAPBSLKH 720
Db 164 II-----PPPPPPPLPPPHITGP 183

RESULT 14

EMIL1_HUMAN ID EMIL1_HUMAN STANDARD; PRT; 1016 AA.
AC Q9Y6C2; Q96G58; Q9UG76;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE EMILIN 1 precursor (Elastin microfibril interface-located protein 1)
DE (Elastin microfibril interface 1).
GN EMILIN1 OR EM1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RX MEDLINE=99287867; PubMed=10358019;
RA Doliana R., Mongiat M., Buccioti F., Giacomello E., Deutzmann R., Volpin D., Bressan G.M., Colombatti A.;
RT "EMILIN, a component of the elastic fiber and a new member of the C1g/tumor necrosis factor superfamily of proteins.";
RL J. Biol. Chem. 274:16773-16781(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20092836; PubMed=10625608;
RA Doliana R., Canton A., Buccioti F., Mongiat M., Bonaudo P., Colombatti A.;
RT "Structure, chromosomal localization, and promoter analysis of the human elastin microfibril interface located protein (EMILIN) gene.";
RL J. Biol. Chem. 275:785-792(2000).
RN [3]
RP SEQUENCE OF 25-1016 FROM N.A.
RC TISSUE=Uterus;
RA Ottenwaelder B., Obermaier B., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 434-1016 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
PL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP TISSUE SPECIFICITY.
RX MEDLINE=21192249; PubMed=11278945;
RA Doliana R., Bot S., Munguerra G., Canton A., Cilli S.P.,
RA Colabattini A.;
RT "Isolation and characterization of EMILIN-2, a new component of the
RT growing EMILINs family and a member of the EMI domain-containing
RT superfamily.";
RL J. Biol. Chem. 276:12003-12011(2001).
CC -!- FUNCTION: May be responsible for anchoring smooth muscle cells to
CC elastic fibers, and may be involved not only in the formation of
CC the elastic fiber, but also in the processes that regulate vessel
CC assembly. Has cell adhesive capacity.
CC -!- SUBUNIT: Homotrimer associated through a moderately stable
CC interaction of the C-terminal globular Clq domains, allowing the
CC nucleation of the triple helix and then a further quaternary
CC assembly to higher order polymers via intermolecular disulfide
CC bonds. Interacts with EMILIN2.
CC -!- SUBCELLULAR LOCATION: Extracellular matrix. Found mainly at the
CC interface between amorphous elastin and microfibrils.
CC -!- TISSUE SPECIFICITY: Distributed in tissues where resilience and
CC elastic recoil are prominent. Highest levels in the adult small
CC intestine, aorta, lung, uterus, and appendix and in the fetal
CC spleen, kidney, lung, and heart; intermediate expression was
CC detected in adult liver, ovary, colon, stomach, lymph node and
CC spleen; adult heart, bladder, prostate, adrenal gland, mammary
CC gland, placenta and kidney showed low expression whereas a series
CC of other adult tissues, including skeletal muscle and different
CC regions of adult brain show no expression.
CC -!- MISCELLANEOUS: Its deposition precedes the appearance of elastin
CC and is simultaneous with that of fibrillin 1.
CC -!- SIMILARITY: Contains 1 Clq domain.
CC -!- SIMILARITY: Contains 1 collagenous domain.
CC -!- SIMILARITY: Contains 1 EMI domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF088916; RAD42161.1; -;
CC EMBL; AF162780; RAP25006.1; -;
CC EMBL; AL050138; CAB43287.1; -;
CC EMBL; BC009947; AAH09947.1; -;
CC PIR; T08772; T08772.
CC Genew; HGNC:19880; EMILIN1.
CC MIM; 130660; -;
CC GO; GO:0005515; F:protein binding; IPI.
CC GO; GO:0007155; P:cell adhesion; TAS.
CC InterPro; IPR001073; Clq.
CC InterPro; IPR008160; Collagen.
CC InterPro; IPR000577; FGGY kin.
CC InterPro; IPR008983; TNF_like.
CC Pfam; PF00386; Clq; 1.
CC Pfam; PF01391; Collagen; 1.
CC Cell adhesion; Extracellular matrix; Signal; Coiled coil; Collagen;
CC Glycoprotein.
KW

FT	SIGNAL	1	21	POTENTIAL.
FT	CHAIN	22	1016	EMI LIN 1.
FT	DOMAIN	57	129	EMI.
FT	DOMAIN	216	256	COILED COIL (POTENTIAL).
FT	DOMAIN	356	420	COILED COIL (POTENTIAL).
FT	DOMAIN	576	603	COILED COIL (POTENTIAL).
FT	DOMAIN	685	752	COILED COIL (POTENTIAL).
FT	DOMAIN	835	857	COILED COIL (POTENTIAL).
FT	DOMAIN	814	864	COLLAGEN.
FT	DOMAIN	872	1007	C10.
FT	CARBOHYD	154	154	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	415	415	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	455	455	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	561	561	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	658	658	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	766	766	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	794	794	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CONFLICT	25	26	PP -> TR (IN REF. 3).
FT	CONFLICT	434	436	GLS -> HEG (IN REF. 4).
FT	SEQUENCE	1016	AA; 106694	NW; 6CF303238DD0EE26 CRC64;
SQ	SEQUENCE	1016	AA; 106694	NW; 6CF303238DD0EE26 CRC64;

Query Match 3.6%; Score 135.5; DB 1; Length 1016;
Best Local Similarity 21.2%; Pred. No. 0.34;
Matches 184; Conservative 71; Mismatches 297; Indels 315; Gaps 43;

QY	47	WRGCRVHRGPD	118	WRCCQGGYDCAESAPALGPASSTPRPLARPAPNLGSSSAGSPSLGGLGEGPGESEK	177
QY <td>47</td> <td>WRGCRVHRGPD</td> <td>118</td> <td>WRCCQGGYDCAESAPALGPASSTPRPLARPAPNLGSSSAGSPSLGGLGEGPGESEK</td> <td>177</td>	47	WRGCRVHRGPD	118	WRCCQGGYDCAESAPALGPASSTPRPLARPAPNLGSSSAGSPSLGGLGEGPGESEK	177
QY <td>86</td> <td>LSQRAHQHQQIDELQWRTG</td> <td>178</td> <td>VSQLEQV-QSLTKELQGLRGLVGLQSLRLAEDVQRAVETAFNCRQPADAAARPGVHET</td> <td>236</td>	86	LSQRAHQHQQIDELQWRTG	178	VSQLEQV-QSLTKELQGLRGLVGLQSLRLAEDVQRAVETAFNCRQPADAAARPGVHET	236
QY <td>127</td> <td>LSYVNSNLQLL</td> <td>177</td> <td>LSYVNSNLQLL</td> <td>177</td>	127	LSYVNSNLQLL	177	LSYVNSNLQLL	177
QY <td>237</td> <td>LNETHQQLLDTRVTHQDLGHLNNHGGSSGSRAPAPAPAPPGPSEELLRLQLEQ</td> <td>296</td> <td>LNETHQQLLDTRVTHQDLGHLNNHGGSSGSRAPAPAPAPPGPSEELLRLQLEQ</td> <td>296</td>	237	LNETHQQLLDTRVTHQDLGHLNNHGGSSGSRAPAPAPAPPGPSEELLRLQLEQ	296	LNETHQQLLDTRVTHQDLGHLNNHGGSSGSRAPAPAPAPPGPSEELLRLQLEQ	296
QY <td>178</td> <td>PLKELISVHF-GEQG</td> <td>297</td> <td>RLQESCVCLAGLDGFRQOQEDRRLRAMEKLLASVEERQRLHAGLAVGRPPPOCCSP</td> <td>356</td>	178	PLKELISVHF-GEQG	297	RLQESCVCLAGLDGFRQOQEDRRLRAMEKLLASVEERQRLHAGLAVGRPPPOCCSP	356
QY <td>214</td> <td>-----NESGLEL-----</td> <td>235</td> <td>-----LTAYNQLCFLDARF-----</td> <td>235</td>	214	-----NESGLEL-----	235	-----LTAYNQLCFLDARF-----	235
QY <td>357</td> <td>ELGRRLAELRLDLVAGSVTVLSGRRTGELGGAAGCGGHPGYTSLASRLSLEDRENS</td> <td>416</td> <td>ELGRRLAELRLDLVAGSVTVLSGRRTGELGGAAGCGGHPGYTSLASRLSLEDRENS</td> <td>416</td>	357	ELGRRLAELRLDLVAGSVTVLSGRRTGELGGAAGCGGHPGYTSLASRLSLEDRENS	416	ELGRRLAELRLDLVAGSVTVLSGRRTGELGGAAGCGGHPGYTSLASRLSLEDRENS	416
QY <td>236</td> <td>-----SLGLFFHWYDSLTGVPAAQALAFKGVNFGNIGALHTQIGARQDR</td> <td>286</td> <td>-----SLGLFFHWYDSLTGVPAAQALAFKGVNFGNIGALHTQIGARQDR</td> <td>286</td>	236	-----SLGLFFHWYDSLTGVPAAQALAFKGVNFGNIGALHTQIGARQDR	286	-----SLGLFFHWYDSLTGVPAAQALAFKGVNFGNIGALHTQIGARQDR	286
QY <td>417</td> <td>TLGPSEEEQESWPAGPGGLSHW</td> <td>465</td> <td>-----LPAARGRL-QLGGLLANVSG-----ELGRLD-</td> <td>465</td>	417	TLGPSEEEQESWPAGPGGLSHW	465	-----LPAARGRL-QLGGLLANVSG-----ELGRLD-	465
QY <td>287</td> <td>CTEGARRAMEAFQRAAGAFSLLENFHSAPDMSAAS--LCALEQLMMAQAQECVPEG</td> <td>344</td> <td>CTEGARRAMEAFQRAAGAFSLLENFHSAPDMSAAS--LCALEQLMMAQAQECVPEG</td> <td>344</td>	287	CTEGARRAMEAFQRAAGAFSLLENFHSAPDMSAAS--LCALEQLMMAQAQECVPEG	344	CTEGARRAMEAFQRAAGAFSLLENFHSAPDMSAAS--LCALEQLMMAQAQECVPEG	344
QY <td>466</td> <td>-----LLE--EQVAGAMQACGLCSGAPGEQDSQVSEILSALERRVL-----</td> <td>505</td> <td>-----LLE--EQVAGAMQACGLCSGAPGEQDSQVSEILSALERRVL-----</td> <td>505</td>	466	-----LLE--EQVAGAMQACGLCSGAPGEQDSQVSEILSALERRVL-----	505	-----LLE--EQVAGAMQACGLCSGAPGEQDSQVSEILSALERRVL-----	505
QY <td>345</td> <td>LSPPASMAPQDCLAQLRL</td> <td>392</td> <td>-----AQEAAQVAAB--YRLVHRTMAQPPVHDVYVVS</td> <td>392</td>	345	LSPPASMAPQDCLAQLRL	392	-----AQEAAQVAAB--YRLVHRTMAQPPVHDVYVVS	392
QY <td>506</td> <td>-----DSEGURLVSGSLHTVEAAGEARQATLEGQVEVGLQDRVDAQDETAAE</td> <td>555</td> <td>-----DSEGURLVSGSLHTVEAAGEARQATLEGQVEVGLQDRVDAQDETAAE</td> <td>555</td>	506	-----DSEGURLVSGSLHTVEAAGEARQATLEGQVEVGLQDRVDAQDETAAE	555	-----DSEGURLVSGSLHTVEAAGEARQATLEGQVEVGLQDRVDAQDETAAE	555
QY <td>393</td> <td>WTALVHVKAEPFRL</td> <td>438</td> <td>-----AHYHVAMALCDGSPATEGEL-----PTHQVFLQPP</td> <td>438</td>	393	WTALVHVKAEPFRL	438	-----AHYHVAMALCDGSPATEGEL-----PTHQVFLQPP	438
QY <td>556</td> <td>FTLRNLTAARLQGLLEGLQAHDGCGACGQVQBELRLDRDGVVERCSCF</td> <td>609</td> <td>FTLRNLTAARLQGLLEGLQAHDGCGACGQVQBELRLDRDGVVERCSCF</td> <td>609</td>	556	FTLRNLTAARLQGLLEGLQAHDGCGACGQVQBELRLDRDGVVERCSCF	609	FTLRNLTAARLQGLLEGLQAHDGCGACGQVQBELRLDRDGVVERCSCF	609
QY <td>439</td> <td>-----TSSKPRGPVLPQELERERQLGKAHLKRAILQCEEARLHALCRVLRVDLLR</td> <td>490</td> <td>-----TSSKPRGPVLPQELERERQLGKAHLKRAILQCEEARLHALCRVLRVDLLR</td> <td>490</td>	439	-----TSSKPRGPVLPQELERERQLGKAHLKRAILQCEEARLHALCRVLRVDLLR	490	-----TSSKPRGPVLPQELERERQLGKAHLKRAILQCEEARLHALCRVLRVDLLR	490
QY <td>610</td> <td>RGFGAGPGVGGPSRGFL</td> <td>654</td> <td>-----DGFSVFGSGSSGSAQAQGLSEVILSF</td> <td>654</td>	610	RGFGAGPGVGGPSRGFL	654	-----DGFSVFGSGSSGSAQAQGLSEVILSF	654
QY <td>491</td> <td>AVISOTLQRLAKYAELEDREDDFCFAAEPDIPKTHQKPEARMPLSGKGPDIPIHRLG</td> <td>550</td> <td>AVISOTLQRLAKYAELEDREDDFCFAAEPDIPKTHQKPEARMPLSGKGPDIPIHRLG</td> <td>550</td>	491	AVISOTLQRLAKYAELEDREDDFCFAAEPDIPKTHQKPEARMPLSGKGPDIPIHRLG	550	AVISOTLQRLAKYAELEDREDDFCFAAEPDIPKTHQKPEARMPLSGKGPDIPIHRLG	550
QY <td>655</td> <td>SLNDSLN</td> <td>679</td> <td>-----ELQTTVEGQGLADLAD-LG</td> <td>679</td>	655	SLNDSLN	679	-----ELQTTVEGQGLADLAD-LG	679
QY <td>551</td> <td>PLS--VFSANRWRLVGPVHLTRGEGGF</td> <td>590</td> <td>-----GLTLRGDSPVL-----IAA</td> <td>590</td>	551	PLS--VFSANRWRLVGPVHLTRGEGGF	590	-----GLTLRGDSPVL-----IAA	590
QY <td>680</td> <td>ATKDRIISINRLQQEATHESEERFRGLBEGQAQACQPSLEGRLGLEGVCEKLDT</td> <td>739</td> <td>ATKDRIISINRLQQEATHESEERFRGLBEGQAQACQPSLEGRLGLEGVCEKLDT</td> <td>739</td>	680	ATKDRIISINRLQQEATHESEERFRGLBEGQAQACQPSLEGRLGLEGVCEKLDT	739	ATKDRIISINRLQQEATHESEERFRGLBEGQAQACQPSLEGRLGLEGVCEKLDT	739

QY	556	SAKRWELVGPVH----	LTRGEGGGLTLRGDSPVLIAAVIPGSQAAAAAGLKEGDYIVSV	611
Db	58	SEESRSIYGLVQRCVIIQDDNGFLTVSGDNPFVQSVKEDGAAMRAGVQTGDRIIKV	117	
QY	612	NGQPCRWWRHAEVVTELKAAGEAGASIQ-----	VVSLLPSSRLPS-LGDRRPVLLGP	662
Db	118	NGTLVTHSNHLEVVKIKSGSYVALTVQGRPPGSPQIPLADSEVEPSPVIGHMSPINTSP	176	

Search completed: July 13, 2004, 12:05:10
Job time : 28 secs

QY	492	VISQTLQSLAKYAEALDREDDFCEAAEAPDIQ-----PKTHOKPEARMPRLSQGKGPDI	546
DB	284	SINGNLDAV-----KENDRVYLMRVPSSSLSPFAFSVMKPMWMTDILDASKENF	336
QY	547	HLRGPLSVFSAKRW-----RLVGFVHLTRGEGFGLTLRG-DSPVLIAAVI	592
DB	337	SILVPDSSAKALRYTEMVDDVIRTAERLQAASELTR-----VRLKEMDLFDSILAV-	389
QY	593	PGSQAAAGLKEGDIYVSVNGQPC-----	616
DB	390	DGNSALPVDLKEDEVAVQISGGPAGLAELOQLRDLKRVNQLLVHTEELLOKEATEDSQ	449
QY	617	-----RWWR-HAEVVT-----LKAAGEAGASLQVSVLLPSSRLPSLGDRR	656
DB	450	FRSQFGTRWTRPQSSLTLYKNLQDLRLNRPANLKAQGESDVKIE-RSVRNSALMSILDRR	508
QY	657	PV 658	
DB	509	PI 510	

RESULT 2
 T37837
 probable signal transduction pathway protein - fission yeast (Schizosaccharomyces pombe)
 C;Species: Schizosaccharomyces pombe
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C;Accession: T37837
 R;Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V. submitted to the EMBL Data Library, September 1997
 A;Reference number: Z21749
 A;Accession: T37837
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-775 <MUR>
 A;Cross-references: EMBL:Z99162; PIDN:CAB16216.1; GSPDB:GN00066; SPDB:SPAC17G6
 A;Experimental source: strain 972h; cosmid c17G6
 C;Genetics:
 A;Gene: SPDB:SPAC17G6.05c
 A;Map position: 1

```

Qy 542 -----GPDIFHRLGLPLSVPSAKN 559
Db 368 AKKIVGDDLFKAFVPSAVITASS 390

RESULT 3
S41034
hypothetical protein R10EL2.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997
C:Accession: S41034
R:Smith, A.
submitted to the EMBL Data Library, January 1994
A:Reference number: S41034
A:Accession: S41034
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-882 <SMI>
A:Cross-references: EMBL:Z29561; NID:g450897; PID:g450898
C:Genetics:
A:Introns: 72/2; 324/3; 359/1; 400/2; 722/1; 743/1; 782/1; 818/1

Query Match 7.5%; Score 284.5; DB 2; Length 882;
Best Local Similarity 22.8%; Pred. No. 1.1e-11;
Matches 132; Conservative 98; Mismatches 209; Indels 139; Gaps 22;

Qy 167 LKETKELDSWSTPLKELIS--VHFGEADGAEAE-IRELEALR-QAMTTP-SRNEGUELL 221
Db 11 LKSTNEVDLVKPLTSYIDNVNTSDNNRSDVAEAVQELNLRKACCOPLDKHQSLADVL 70

Qy 222 TAYYNQLCFDLARELITPARSLGLFFHWYDSL-TGVPAQORAL-----APEKGSVLENI 274
Db 71 TRYDQILVAIENKIIISATQNPVVKWDAPDKGSLFSRASLSLDSGSPERAVALFNIG 130

Qy 275 ALHTQIGARODRSCTEGARAMEAFORAAAGFSLRLRE---NFSHAPS.PDMSAASLCALE 330
Db 131 SLMSQIGAAQOFHTDDIEIKYSAKL.FQSAGVFAFLRDVLVGMWQOEPTPDLMPDTLAALS 190

Qy 331 QLMMAQAEQCEVFEGLSPSPASMAPODCLAQLRLAQEAACVAAYKLVHRTMAQPPVHDVVP 390
Db 191 ALMTAQAEAIY-----IKGHDKMKATSMVKISIAQVAEFYSEAOKMMSKDIVRGLWD 243

Qy 391 VSWTALVHVKAEYFRSLAHYHVAVALCDGSPATGELP.THQVFLQPTSSKPRGPVL.PQ 450
Db 244 KDWSAIYSGKNLAYCALAQYHQ$-EVCG-----

Qy 451 ELEFRRLQGAHLKRAILGQBEALRLHALCRLVREVLLRAVISQ-----TLQSLAK- 503
Db 271 ---EARQIGE-----QLSRLAESLKLFTAQYKLPDITGIWDIYPSVSKA 313

Qy 504 YAELEDREDDFCEAAEAPDIQKTHQKPEARMPRLISQQKGP-----DIFHRLGLPLSVPSA 557
Db 314 HAAAKKDNDFIYHEKVSDFR-TLPTLPKAVLAKETPMQTMPTSPFRDMFAVLVPVQVHNA 372

Qy 558 KNRW-----RLVCPVHLTRCEGGFGILTRGDSVPLTAAVIPGSAQAAGLKE 604
Db 373 MCQSYDARKAEVLNMTVMREATQLMNG-----VLASLNLPAL----- 411

Qy 605 GDYIVSVNGQPCRWWRHAEVVTLEKAA--GEAGASLOWVSLPSRLPSLDGRRRPVLIGP 662
Db 412 -DDVTSTETLP-----ESLKLSAKLXKNGSSSEIMRLF--SELP.TLYQRNEDIILTE 460

Qy 663 RGLLRSQREHCKTTPASTWASPRLLNWSRKAQOQKGTG 700
Db 461 T$RIILNEKESD$D$TWRKOLG$-----KWTRMSS$QLTG 493

```

153 FIFVZLDDANWISFYIKVMILIAQAQECVLQKRSLLDNKSAI VI AKLSLWLEQAYDSAARIKD 212

344 --GLSPASMAPQDCIAQLRLAQEAQAQVAEYRLVHRTMAQPFVHDVYPVSMTALVHVK 401

QY 344 --GLSPASMAPQDCLAQLFLAQEAAQVAAAEYRLVHRTMAQPPVPHDYPVPSWTALVHVKA 401

Db 213 EWSNIPESVQRITAIKLDLKQNAEKENDP-----IYHDRVPKQEDAIQVOK 262
QY 402 E-----YFRSLAHVHVMALCDGSPATEGELPHEQVFLPPTSSKRGPGVLPQLEER 455
Db 263 DGGAMCKVKTSLSPDLDPVSC-GCDLFGKLLPTFVQDAVKYSDK-----DQALRE- 314
QY 456 ROLGAHLKRAILGOEEALRLHALCRVLRREVDLLRAVISOTLQSLAKYAEILDREDDFCE 515
Db 315 -----IKCVRSVDEHLNVLQ-----QAEFDKURFMLENEG-KRSREAWFELS-EDLMKR 362
QY 516 AAEA---PDIQPKTHQK---PEARMPRLSQGKGPDI FHRGLPSVFSAKNRWRLVGPVHLT 570
Db 363 NADWTSYEDCVNLDKMGESSDTRVAEAKINTLLSKLRAIDLQKSS-----I 411
QY 571 RGEFGFGLTRGDSPLVIAAIPVPSQAAAAGLKEGD-YIVSVNGQPCRWNRHAEVVVTLK 629
Db 412 RSDEGF-----ILIQKEL-----ERLAHLEQAKAHNVSLN-----K 443
QY 630 AAGRAGASLQVVSL 643
Db 444 AIAQHSANLQLTL 457
RESULT 7
JC2129
protein kinase PKN (EC 2.7.1.-) - human
N:Alternate names: protein-kinase-C-related kinase PRK1
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1993 #sequence revision 20-Aug-1994 #text_change 17-Nov-2000
C:Accession: JC2129; S67526; S51020; I53327
R:Mukai, H.; Ono, Y.
Biochem. Biophys. Res. Commun. 199, 897-904, 1994
A:Title: A novel protein kinase with leucine zipper-like sequences: Its catalytic domain
A:Reference number: JC2129; MUID:94183274; PMID:8135837
A:Accession: JC2129
A:Molecule type: mRNA
A:Residues: 1-942 <MUK>
A:Cross-references: GB:D26181; NID:9473712; PIDN:BA05169.1; PID:9825505
R:Palmer, R.H.; Riddien, J.; Parker, P.J.
Eur. J. Biochem. 227, 344-351, 1995
A:Title: Cloning and expression patterns of two members of a novel protein-kinase-C-rela
A:Reference number: I53327; MUID:95154310; PMID:7851406
A:Accession: S67526
A:Molecule type: mRNA
A:Residues: 1-190,'D',192-942 <PAL>
A:Cross-references: EMBL:S75546; NID:9914097; PIDN:AAB33345.1; PID:9914098
R:Palmer, R.H.; Riddien, J.; Parker, P.J.
FEBS Lett. 356, 5-8, 1994
A:Title: Identification of multiple, novel, protein kinase C-related gene products.
A:Reference number: S51020; MUID:95080426; PMID:7988719
A:Accession: S51020
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: mRNA
A:Residues: 700-799,'A' <PA2>
C:Comment: This protein has a protein kinase domain related to protein kinase C.
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C:Keywords: ATP; leucine zipper; phosphotransferase; protein kinase
F:39-66/Region: basic
F:70-287/Region: leucine zipper motif
F:613-874/Domain: protein kinase homology <KIN>
F:621-629/Region: protein kinase ATP-binding motif
F:644/Active site: Lys #status predicted
Query Match 4.1%; Score 155.5; DB 2; Length 942;
Best Local Similarity 21.4%; Pred. No. 0.012;
Matches 159; Conservative 76; Mismatches 260; Indels 247; Gaps 36;
QY 69 QLHAVGDSILTIQCQLOSRRAQIHQIQDKELQMETGAENLYRATSN-NRVRETVALEL 127
Db 18 QLGLAGADLAAPGVQQLERERURERREIKELKKEGAEENLRATYDLSRSLGPVELLL 77

QY 128 SYVNSNQLLKBELBELSGGV---DP-GRHGEAVTVPMIPGLKTKETKELDMSTPLKELI 183
Db 78 RGSRRLLDHLHQQLQELHAAHVLPDPAATHDG-----PQSP----- 113
QY 184 SVHFGEDGASVEA-EIRELEALRQAMRTSPSNESGLELLTAYYNQLCFLDARFLTPARSL 242
Db 114 ---GAGGPTCSATNLSRVAGLEKOLAELEKVKQGAENMIQTYNSGTCKDRKLLTAAQOM 169
QY 243 GLFFHWDYSLT-----GVPAQRALAFEKGSVLFNIGALHTQIGA----- 282
Db 170 -----LQDSKTKIDIIRMQLRRALQAGLENAQAPDDTQSGP--DLGAVELRIELRHHF 222
QY 283 RQDRSCTEGARRAMEAFQRAAGAFSLLRNFSPHAPSMDMSAASLCALQLMMAQAQECVF 342
Db 223 RVEHAVAEGAKNVLRL-----LSAAKAPDRKAVSEAAQEKLTSNQKLGLLR 268
QY 343 EGLSPPSMAPODCLAQLRLAQEAQAQVAAEYLVHRTMAQPPVHDY-----VPVSWTALVH 398
Db 269 EALERRIGELPAD-HPKGRLLREELAAASSAFTRLAGPPFATHYSTLCKPAPLTGTILE 327
QY 399 VKAEYFRSLAH---YHVAMALCDGSPATEGELPHEQVFLQPP----- 438
Db 328 VRVVCRRDLPETIPWNTPSM--GGPGT-----PDSRPFFLSRPARGLYSRSGSLSGRSSL 381
QY 439 -----TSSKPRGP-----VLPQELERRQLGKAHLKRAILGQ 470
Db 382 KAEAEENTSEVSTVLKLDNTVVGTQTSWKPCGNAMDQSFLELERARE-----LELAVFWR 436
QY 471 BEALRLHALC--RVLRVDLLRAVISOTLQSLAKYAEILDREDDFCEAAEAPDIQPKTHQ 528
Db 437 DQ---RGLCALFKLKEDEL-----DNERHEVQDMEPQGCCLVAEV-----TFR 477
QY 529 KPE-ARMPRL-----SQCKGPDIFHRLGPLSVFSAKNRWRLVGPVHLTR----- 571
Db 478 NPVIERIFLRQKKIFSKQQGKA---FQARQNMIDVA--TW-----VRLRLRLIPNAT 527
QY 572 GEGFGFLTRGDSPLVIAAIPVPSQAAAAGLKEGDIYIVSVNGQPCRWNRHAEVVTELKAA 531
Db 528 GTGTG-----SP---GASFGSEARTTG-----DISVE-----KLNIIG 556
QY 632 GEAGASLQVSVLLPSSRLPSLGDPRVLLGPRGLLRSQREHCKCTKPASTWASPRLLNWS 691
Db 557 TDSDSPOKSRDRPSPSSSSLS--FI-----QESTAPELPSETQETGPGAL--- 601
QY 692 RKAQOQKGTGCTPCPCAPVKPAP 713
Db 602 -----CSPLRKSP 609
RESULT 8
I46532
protein co-factor - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 05-Nov-1999
C:Accession: I46532
R:Weinman, E.J.; Stepien, D.; Wang, Y.; Shenolikar, S.
J. Clin. Invest. 95, 2143-2149, 1995
A:Title: Characterization of a protein cofactor that mediates protein kinase A regulation
A:Reference number: I46532; MUID:95256437; PMID:7738182
A:Accession: I46532
A:Status: preliminary; translated from GB/EMBL/DOBJ
A:Molecule type: mRNA
A:Residues: 1-358 <WEI>
A:Cross-references: EMBL:U19815; NID:9687674; PIDN:AAA80218.1; PID:g687675
A:Superfamily: GLGF domain homology
F:19-90/Domain: GLGF domain homology <GLG2>
Query Match 3.7%; Score 141.5; DB 2; Length 358;
Best Local Similarity 24.0%; Pred. No. 0.028;
Matches 89; Conservative 31; Mismatches 96; Indels 155; Gaps 16;
QY 419 GSPATG-----ELPTHEQVFLQPPETSSKPRGP-----VLPQELRE 454

Db 45 GSPAEGAGLLAGDLRVBNGENVKETHQV-----SRRAALNAVRLVVDPDTDEQ 98
QY 455 RROLG---KAHLKRAILGQEARLHLALCRVLREVDLLRAVISQTLQSLAKYAELDRED 511
Db 99 FRKLGVIQIRGELLRAQAPGAPPA-----126
QY 512 DFCEAAEAPDIQKTHQK--PEAR--MPRLSQGKGPDI FHLRGLPLSVFSAKNRWLVGPV 567
Db 127 --GEOGAGENEPEVEKSHPERELRPLCAMK-----158
QY 568 HLTRGEGGGLTURGDS--VJIAAVIPGSAAGLKEGDYIVSYNGQPCRWWRHAEV 624
Db 159 ---KGPNGYGNLHSDKSRPGQFIRAVDPDS PAEASGLRQDRIVEVNGVCVEGKHGDV 215
QY 625 VTELKAGEAGASIQV-----SLLPSR-----LP-----SLGDRRPVLLGP 662
Db 216 VTAIKAGGDE-AKLLVVVDKETDEFFKCKVVPSEHLNGPLPFPFTNGEIQKNPETLAP 274
QY 663 R-----GLLR-----SOREHGCKTPASTWASPRPLNWS-----RK 693
Db 275 AAESPRPALARSASSDTSEELASQDSPKEDSTAPSTSSSDPILDFISLAVAKERA 334
QY 694 AQQCKTGGCPQ 704
Db 335 HOKRSSRRAPQ 345

RESULT 9

A54602

microtubule-associated serine/threonine protein kinase MAST205 - mouse

C:Species: Mus musculus (house mouse)

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999

C:Accession: A54602

R:Walden, P.D.; Cowan, N.J.

Mol Cell. Biol. 13, 7625-7635, 1993

A:Title: A novel 205-kilodalton testis-specific serine/threonine protein kinase associated

A:Reference number: A54602; MUID:94067123; PMID:8246979

A:Accession: A54602

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1734 <NAL>

A:Cross-references: GB:U03313; NID:9406057; PIDN:AAC04312.1; PID:9406058

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

C:Keywords: ATP

F:451-726/Domain: protein kinase homology <KIN>

F:459-467/Region: protein kinase ATP-binding motif

Query Match 3.7%; Score 141; DB 2; Length 1734;
Best Local Similarity 21.3%; Pred. No. 0.28;
Matches 135; Conservative 62; Mismatches 238; Indels 200; Gaps 26;

QY 161 PMIPGLKETKE-----LDWSTPLX---ELISVHFGEGCASY-----194
Db 708 PLERLGTSSAYVKQHPFFMGDLDTGLLRQKAEFIPQLESEDDTSYFDRSRYYHHVDS 767
QY 195 -BAEIRE---LEALROAMRTP--SRNESGLELLTAYYNOLCFDARFLPARSLGIFHH 248
Db 768 DEEVESEDCLEIRQFSSCSPPFSKYSSNE-----RLSLEER-RTP-----809
QY 249 YDSLGTGVPQAQALAFKSGVLFNIGALHTQIGARQDRSCTEGARRAMEAFQAAAGFSL 308
Db 810 -----PPTKESLSSEKEDHSDGLAGL-----KGRDRSWIGVSPPELLR--XELSVSESS 855
QY 309 LRENFSHAPSMDMSAASLALQLMWAQAQCVFEGLSPPASMAPQDCIAQLRLAQEAQA 368
Db 856 HTESDS---SPMTVRHRCs-----GLPDGPHCFEETSSTP-----888
QY 369 VAAEYRLVHRTMAQPPVHDYVVPVSWTALVHVVAEYFRSLAHYHVAMALCDGSPATEGELP 428
Db 889 -----RKQQQEGIWLIIPSG-----EGSSRPVPERP 915
QY 429 THEQVFL--QPPTSSKPRGVLFOBLEERRQLGKAHLKAILGQBEALRLHALC-----480

Db 916 LERQLKLDDEPPQGSQSCCPALETRGRTQPLABEATAKAISDLAVRRARHRLSGDSIE 975
QY 481 -RVLRVEDLLRAVISQTLQSLAKYAELDREDDFCEAAEAPDIQPKTHQKPEAR-----533
Db 976 KRITRPVNVYIKASATALSLL-----IPSEHACSPASPMSPHSOSSPSSDRSLSP 1030
QY 534 --MPRLSQGKGPDI FHLRGLPLSVFSAKNRWLVGPVHLTRGEGGFGTLR-----GDSP 585
Db 1031 DFLPALGSLRPPIIIRHAGK-----KYGFTLRAIRVYMGDTD 1067
QY 586 V-----LIIAAVIPGSAAGLKEGDYIVSYNGQPCRWWRHAEVVTTELKAGEAGASLQV 640
Db 1068 VYVTHMVHVEHVGEGPASEAGLRQGLITHVNGEPVHGLVHTVEVVELVLSGNGK-VSIST 1126
QY 641 VSLPSRLPSLQDRRPVLLGPRGLLRQSREHG-----CKTPASTWASPR 685
Db 1127 TPLENTS--IKVGPARKGSYKAKWARRSKSKGDKQESKRSLFKKITQASLLHTSR 1184
QY 686 PLLNWRKKAQKQKGTG-GCPQPCAPVKP-APPSSLK 718
Db 1185 SLSSLNRLSLSGSGSPGSPTHSHLSLSPRPPOGYR 1219

RESULT 10

T34847

probable transcription regulator - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999

C:Accession: T34847

R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, February 1999

A:Reference number: Z21559

A:Accession: T34847

A>Status: preliminary; translated from GB/EMBL/DDBT

A:Molecule type: DNA

A:Residues: 1-943 <OLI>

A:Cross-references: EMBL:AL035478; PIDN:CAE36601.1; GSPDB:GN00070; SCOEDB:SC2G5.14C

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SC2G5.14C

Query Match 3.7%; Score 139.5; DB 2; Length 943;
Best Local Similarity 22.3%; Pred. No. 0.15;
Matches 177; Conservative 62; Mismatches 253; Indels 301; Gaps 39;

QY 55 GEDPVRV-----HVGPMNQLHAGVCDLSLTQICGQLQSPR--AQIH 94
Db 166 GETPVRLLCVTEGTETDDAEYDHLRASPPDTRAVRLGPLTRAQVSAALDGRGVSGUH 225
QY 95 QQIDKELQMRGAEINLYRATSNRRVETVAL-----ELSYVNSNLQLLKEELELSG---146
Db 226 RSTVRDIHRTSGGNPLFALELGRALAEFPAPRPGEPVPVTSRLALVLSRLEMLSGEAR 285
QY 147 -----GVDP-----GRHGEANTVPMPIGL-----KETELDWSTPLKELISVH 186
Db 286 RYLLVASAGARTPALLRAAGRENAEETVQAALGLLAPDPEGPTLRFAPH--LIS---340
QY 187 FGEDGASY-EAERIRLEALROAMRTPSNESGLELLTAYYNOLCFDARFLPARSLGIF 245
Db 341 ---AALYAEAPAPERAAHAAALSTAADP-----IERARHLAL- 375
QY 246 FHWYDSLTVGP-----AQRALAFKSG--SVLFNIGAL---HT-----QIGA 282
Db 376 -----ATTGTDPEVAARLAEAAALARDGAPSAASLGLLAARHTPADGTPGPDERRUTA 430
QY 283 ROD-----RSCTEGAR-RAMEAFORAGA-----305
Db 431 AEDAITAGEQDLARDVAREVLTRATAPGERVRAWMVAITEAQAQATGEVDVFPQALADAG 490
QY 306 -----FSLLRNESHAPSMDMSAASLCA-----LEOLMNA- 335
Db 491 DDPRLALVHYQLAWSRLVVGDFQAQREAAH-----AARLAARAGERSTELLALAF 543


```
Db      297  SGLHLSPTAAEARRLEPCBSTARHRWGTGTVKSSATSEPLPACILGTGLGPIPHGPWAS 356
QY      684  PRPILNWSRAQOQKTCGCPQCAPVVPKPA---PPSSLKHFG 721
Db      357  ACPEL-----PQFWTGGWSCHCFEISPSGPEPPSCPCPPG 392

RESULT 13
T08772
hypothetical protein DKF2p586M121.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 13-Aug-1999
C;Accession: T08772
R;Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, May 1999
A;Reference number: Z16474
A;Accession: T08772
A;Molecule type: mRNA
A;Residues: 1-992 <OTT>
A;Cross-references: EMBL:AL050138
A;Experimental source: adult uterus; clone DKF2p586M121
C;Genetics:
A;Note: DKF2p586M121.1

Query Match      3.6%; Score 135.5; DB 2; Length 992;
Best Local Similarity 21.2%; Pred. No. 0.31;
Matches 184; Conservative 71; Mismatches 297; Indels 315; Gaps 43;

QY      47  WRGCRVHRGED-----PVRVHVGPMNPQLHVAVGDS-LTQI-----CCQG 85
Db      94  WRCCOYGGDDCAESAPALGPASSTPRLPRLARPENLSSSAGSLGGLGGPGCESEK 153
QY      86  LQSRRAQHQQIDKEYQMRTG-----AENLYRATS---NNR-----VRETVALE 126
Db      154  VQOLEEQV-QSUTKELQGLRGVLQGLSRLAEDVQAVETAFNGRQQPADAAARPGVHET 212
QY      127  LSYVNSNLQLL-----KEELEELSG--GVDPGRHGSEAVTPMIPLGKTKETKELDWST 177
Db      213  LNEIQHQLQLLDFRVSTHQELGHNNHGGSSSGSRAPAPASAPPGPSEELLKQLEQ 272
QY      178  PLKELLSVHF-GBDG-----ASYAEIREFALROAMRTPSR-----213
Db      273  RLQESCSVLAGLDGFRQOQEDRERLRAMEKLLASVEERQRLHLAGVGRPPQCCSP 332
QY      214  -----NESGLEL-----LTAYYNQLCFELDARF--235
Db      333  ELGRRLAELERLDVVAGSVTVLSGRRGTELGGAQGGHPGYTSLASRLSRLDEDFNS 392
QY      236  -LTPAR-----SLGLFFHWYDSLITGVPAAQORALAFKSGVLEFNGALHTQIGARQDR 286
Db      393  TLGPSEEQEESWFGAPGGLSHW-----LPAARGRLLE-QLGGLLANVSG--ELGRLD- 441
QY      287  SCTEGARRAMEAFQRAAGAFSLLENFNSHAPSPDMSAAS--LCALFQLMMAQAQECVFE 344
Db      442  -----LLE--EQVAGAMQACGLQCSGAPGEQDSQVSEILLSALERRVL-----481
QY      345  LSPSPASMAPQDCIQLRL-----AQEAQAQVAE--YRLVHRTMAQPPVHDVVPVS 392
Db      482  -----DSEGQRLVGSGLHTVEAAGEARQATLEGLQEVVGVRLQDRVDAQDETAAE 531
QY      393  WTALVHVKAERYFRSL-----AHVHVAMALCDGSPATEGEL-----PTHEQVFLQPP 438
Db      532  FTLRNLNLTAARLCOLEGLLQAHGDECGACGGVQBELGRLRDGVERCSCP-----LLPP 585
QY      439  -----TSSPRGPVLPQLEERLQGLKAHLKRAILQGEAEALRIHALCRVIREVDLLR 490
Db      586  RGPAGPGVGGPSRGPL-----DGFSVFGSGSALQALQGLSEVILSF 630
QY      491  AVISQTLQSLAKYAELEDREDDFCENAEAPDIQPKTHQKPEARMPLSQGKGPDIPIRLG 550
Db      631  SSINDSLN-----ELQTTVEGQADLAD-LG 655
QY      551  PLS--VFSAKNRWRLVGPVHLTRGEGGF-----GLTLRGDSPVL-----INA 590
```

```
Db      656  ATKDRIISEINRLQOEATEHATESERFRGLERGOAQGCPSLEGRGLRGLEGVCERLDT 715
QY      591  VIPGSAAMAAGLKEGDYIVSVNGQPCRW---RHAENVVTELKAA-----GEAGASLQV 640
Db      716  VAGGLQ-----GLREG-----LSRHVAGLWAGLRETNTTSQQAALLKLVGGQAGLGRRL 766
QY      641  VSLPSSRLPSLGRD-----RPVLLGFRGLLRQSRHEGCKTTPASTWASPRPLLNWSRKAQ 695
Db      767  GAL--NSSLQLLEDRLHLQLSLXDLTGPA-----EAGPPGPPGLQGPFGPAGPPGSPGK 818
QY      696  QGKTG--GCPQP-----CAPVKPAP 713
Db      819  DQEGPIGPFPQPGQEQGVGEGAPAAPVP 845

RESULT 14
T09144
probable guanine nucleotide exchange factor RhoGEF2 - fruit fly (Drosophila melanogaster)
N;Alternate names: Shar pei/DRHOGEF2
C;Species: Drosophila melanogaster
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 08-Sep-2002
C;Accession: T09144; T09223
R;Haecker, U.; Perrimon, N.
submitted to the EMBL Data Library, October 1997
A;Reference number: Z16586
A;Accession: T09144
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2559 <HAE>
A;Cross-references: EMBL:AF031930; NID:g2687355; PID:g2687356
R;Barrett, K.; Leptin, M.; Settlement, J.
Cell 91, 905-915, 1997
A;Title: The Rho GTPase and a putative RhoGEF mediate a signaling pathway for the cell st
A;Reference number: Z16618; MUID:98088790; PMID:9428514
A;Accession: T09223
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-676,'L',678-837,'L',839-889,'S',891-1243,'D',1245-1358,'E',1360-1368,'R',13
A;Cross-references: EMBL:AF032870; NID:g2760367; PIDN:AAC38820.1; PID:g2760368
C;Genetics:
A;Gene: rhoGEF2
A;Cross-references: FlyBase:FBgn0023172
A;Map position: 2; 53F1-2
A;Note: orchestrates cell shape changes during gastrulation
C;Function:
A;Description: mediates actin rearrangements required for cell shape changes during gastr
C;Superfamily: rat Munc13-3 protein; protein kinase C zinc-binding repeat homology
C;Keywords: embryo; GTP exchange; signal transduction
F;1151-1200/Domain: protein kinase C zinc-binding repeat homology <KZN>

Query Match      3.6%; Score 135.5; DB 2; Length 2559;
Best Local Similarity 25.1%; Pred. No. 1.2;
Matches 50; Conservative 27; Mismatches 71; Indels 51; Gaps 6;

QY      483  LREVDLLRAVISQTIORSIAKYAEILDRE--DDFCFAAEAPDIQPKTHQKPEARMPLRSQ 540
Db      202  LNQSDNSNPVLQAPGERSNLNLTPLSRDLSGGHTQBSTTPTATPTST---PSLALPKNFQ- 257
QY      541  KGPDIFHRLGPLSVFSAKNRWFLVGPVHLT-----RGEQGFGLTLRGDSPVLIAAVIPG 596
Db      258  -----YLTLTVRKDSNGYGMKVGSDNPFVFEVSKPGGA 290
QY      597  AAAAGLKEGDYIVSVNGQPCRWHRHAENVVTELKAAEA-----GASLQVSVLL 644
Db      291  RELIAGVAGDMLLRVNGHEVRLEKPTVVGLIKASTVELAVKRSQKLTRPSSSVVT-- 348
QY      645  PSSRLPSLGRRRPVLGLGR 663
Db      349  ESTPILSGRDRTASITGPQ 367

RESULT 15
```

T14355
protein-tyrosine-phosphatase (EC 3.1.3.48) TD14 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T14355
R;Cao, L.; Zhang, L.; Ruiz-Lozano, P.; Yang, Q.; Chien, K.R.; Graham, R.M.; Zhou, M.
J. Biol. Chem. 273, 21077-21083, 1998
A;Title: A novel putative protein-tyrosine phosphatase contains a BRO1-like domain and s
A;Reference number: 218004; MUID:96361981; PMID:9694860
A;Accession: T14355
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1494 <CAO>
A;Cross-references: EMBL:AF077000; NID:g3598973; PID:g3598974; PIDN:AAC62959.1
A;Experimental source: brain
C;Genetics:
A;Gene: PTP-TD14
C;Function:
A;Description: may be involved in regulating Ha-ras-dependent cell growth
C;Keywords: phosphoric monoester hydrolase

Query Match 3.5%; Score 130.5; DB 2; Length 1494;
Best Local Similarity 23.5%; Pred. NO. 1.2;
Matches 56; Conservative 36; Mismatches 87; Indels 59; Gaps 9;

QY 333 MMAQAQECVFEGLSPPAGMADQCLALQLRLAQEAQAAVAAEYRLVHRTMAQPPVHDY---V 389
Db : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
1 MLGQAQECLE-----KSMLE--DNEKSLVLAISAQVVVYKACRALENPTASLLGRI 53

QY 390 PVSMTALVHVKAERYFRSLAHYHVMALCDGSPATEGELPTHEQVFLQPPTSSKRGPVLP 449
Db ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
54 QKDWKKLVQMKIYFPAVAHLHMG----- 77

QY 450 QELEERRQLGK--AHLKEALIGQEEALRLHALCRVLRVDLLRAVISQTLQESLAKYAEEL 507
Db :: ||:: : ||:: : ||:: : ||:: : ||:: : ||:: : ||:: : ||:: :
78 KOAEQQKFGERVAYFQSGALDKLNEAIKL-----AKGQDPTVQDRLRFAMDVIGGKYN 132

QY 508 DREDDFC--EAAEAPD-IQPKTHQ---KPEARMRLSQGKGPDI FHRIGPLSVFSAKN 559
Db ::||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
133 KQNDFIYHEAVPALDTLPVKGAFLVKPLPVNPTDPAVTGPDIFAKLVFMAAHEASS 190

Search completed: July 13, 2004, 12:07:49
Job time : 37 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2004, 08:21:11 ; Search time 87 Seconds
(without alignments)

2348.067 Million cell updates/sec

Title: US-10-697-266-2

Perfect score: 3774

Sequence: 1 MILLERPDGAGAESPRLQ.....QPCAPVKPAPPSLKHPGW 723

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001s:*

5: Geneseq2002s:*

6: Geneseq2003as:*

7: Geneseq2003bs:*

8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3774	100.0	723	5	AAO18602 Human pro
2	3449.5	91.4	670	6	AB999400 Amino aci
3	2331	61.8	634	5	AAO18603 Human pro
4	2331	61.8	643	6	ABM32685 Human Rho
5	1401	37.1	272	6	ABR41761 Human DIT
6	1278	33.9	257	3	ABR43571 Human can
7	1278	33.9	257	3	AAU00869 Human can
8	1255	33.3	611	5	ABU65216 Human NOV
9	1255	33.3	686	4	AAU10192 Human pro
10	1249	33.1	686	5	ABG94709 Human GTP
11	1236	32.8	666	6	ABG68561 Human GTP
12	1195	31.7	686	5	ABU65215 Human NOV
13	1107	26.7	1111	4	ABG08369 Human NOV
14	972	25.8	556	5	ABG61936 Novel hum
15	817	21.6	648	4	ABG63879 Prostate
16	766	20.3	488	4	ABG64579 Human tra
17	590.5	15.6	321	4	ABG06219 Novel hum
18	413.5	11.0	110	3	ABU41991 Human ORF
19	394.5	10.5	311	4	ABG08364 Novel hum
20	380.5	10.1	939	6	AD454327 Human pro
21	380.5	10.1	1636	5	AAE25081 Human kin
22	377.5	10.0	1636	3	ABG29661 Human his
23	362.5	9.6	308	4	ABG06220 Novel hum
24	341	9.0	170	3	AAE41575 Human ORF
25	304.5	8.1	1583	3	AAE29662 Human tyr

ALIGNMENTS									
RESULT 1									
AAO18602									
ID	AAO18602	standard; protein; 723 AA.	868	3	AA41716	868	3	AAU28082	Novel hum
XX	AAO18602;		7.3		AAU28082	7.3		AB57450	Human sec
XX			7.3		AB57450	7.2		AB90656	Human sec
DT	24-OCT-2002	(first entry)	7.1		AB90656	7.1		ABP41883	Human ova
XX			7.1		ABP41883	7.1		AB57451	Human sec
DE	Human protein kinase N family kinase.		7.0		AB57451	7.0		AB90568	Human sec
XX			7.0		AB90568	7.0		ABG65430	Human alb
KW	Human; protein kinase N; kinase; PKN; eye retinoblastoma; bocio tumour;		7.0		ABG65430	7.0		AB71739	Drosophil
KW	placenta choriocarcinoma; germ cell; leukocyte; uterus tumour; cancer;		6.3		AB71739	6.3		AB67517	Drosophil
KW	pre-B cell acute lymphoblastic leukaemia; Wilm's tumour; enzyme;		5.9		AB67517	5.9		AA97156	Aspergill
KW	brain anaplastic oligodendroma; uterus endometrial adenocarcinoma;		5.1		AA97156	5.1		AA40651	Human ORF
KW	cytostatic; gene therapy; chromosome 8.		4.9		AA40651	4.9		ADD12509	PDZ ligan
XX	Homo sapiens.		4.4		ADD12509	4.4		AB64827	Drosophil
OS			4.4		AB64827	4.4		AB58814	Drosophil
XX			4.3		AB58814	4.3		ABE08492	Novel pro
FN	WO200261062-A2.		4.1		ABE08492	4.1		ABG08365	Novel hum
XX			4.1		ABG08365	4.1		AD538413	Human pro
PD	08-AUG-2002.		4.1		AD538413	4.1		AAW56699	Modified
XX			4.1		AAW56699	4.1		AD56061	Human Pro
PF	29-JAN-2002; 2002WO-US002152.								
XX									
PR	01-FEB-2001; 2001US-00773371.								
XX									
PR	07-MAY-2001; 2001US-00849334.								
XX									
PA	(PEKE) PE CORP NY.								
XX									
PI	Rusch D, Ketchum KA, Di Francesco V, Beasley EM;								
XX									
DR	WPI; 2002-608516/65.								
XX									
DR	N-PSDB; AAL48837, AAL48838.								
XX									
PT	New human kinase peptide and nucleic acid molecule, useful for treating								
PT	disorders associated with abnormal expression of kinase protein, e.g.								
PT	retinoblastomas, Wilm's tumor, in drug screening assays and								
PT	pharmacogenomic analysis.								
XX									
PS	Claim 1; Fig 2; 76pp; English.								
XX									
CC	The present invention provides the protein and coding sequences of a								
CC	novel human protein kinase N family protein. The sequences can be used in								
CC	the treatment of disorders associated with the absence of, inappropriate,								
CC	or unwanted expression of the protein, e.g. eye retinoblastoma, Wilm's								
CC	tumour, placenta choriocarcinomas, bocio tumours, pre-B cell acute								
CC	lymphoblastic leukaemias, uterus tumours, brain anaplastic oligodendromas								
CC	and uterus endometrial adenocarcinomas. The present sequence is the								

DR N-PSDB; ABV72511.
 XX INTSIG polypeptides, useful for diagnosing, preventing, and treating
 PT disorders associated with an abnormal expression or activity of INTSIG,
 PT e.g. neuromuscular, immunological, cardiovascular disorders, cancer and
 PT infection.
 XX
 PS Claim 1; Page 105-106; 126pp; English.
 XX
 CC The present sequence represents an intracellular signalling molecule.
 CC Such molecules are designated INTSIG. The INTSIG polypeptide or its
 CC fragments, and the INTSIG polynucleotide are useful in diagnosing,
 CC preventing, and treating disorders associated with an abnormal expression
 CC or activity of INTSIG, such as neurodegenerative disorders (e.g.
 CC Parkinson's disease, Alzheimer's disease), muscular disorders (e.g.
 CC myotonic dystrophy, catatonias), endocrine disorders (e.g. diabetes,
 CC Grave's disease), cancers (e.g. leukemia, cervical or breast cancers),
 CC immunological disorders (e.g. scleroderma, systemic lupus erythematosus,
 CC allergies), gastrointestinal disorders (e.g. Crohn's disease), renal
 CC disorders (e.g. Goodpasture's syndrome), infections (e.g. viral,
 CC bacterial, fungal, parasitic, protozoal, helminthic), cardiovascular
 CC disorders (e.g. atherosclerosis), or hepatic diseases (e.g. cirrhosis).
 CC INTSIG or its fragments may also be used in screening for compounds that
 CC specifically bind to and modulate the activity of INTSIG. The
 CC polynucleotides can be used to create humanized animals or transgenic
 CC animals to model human disease
 XX
 SQ Sequence 670 AA;

Query Match 91.4%; Score 3449.5; DB 6; Length 670;
 Best Local Similarity 92.7%; Pred. No. 0;
 Matches 670; Conservative 0; Mismatches 0; Indels 53; Gaps 1;
 1 MILERPPGAGAGESPRLOIQRKPRKTRVSSLRGRREGLRDVCAMRCVRHGEDPVR 60
 1 MILERPPGAGAGESPRLOI----- 20
 61 VHVGPMPQLHVGCDLSLQIOCGLOQSPRAOIHQIDKELQWRTGAENLYRATSNRVR 120
 21 -----GCDLSLQIOCGLOQSPRAOIHQIDKELQWRTGAENLYRATSNRVR 67
 121 ETVALELSVNSNQLQKLEBELSGVDPGHSGEAVTPMPLGLKETKELDWSTPLK 180
 68 ETVALELSVNSNQLQKLEBELSGVDPGHSGEAVTPMPLGLKETKELDWSTPLK 127
 181 ELISVHFGEDGASYAEIRELEALRQAMRTPSRNESGLELLTAYYNQLCFLDARFLTPAR 240
 128 ELISVHFGEDGASYAEIRELEALRQAMRTPSRNESGLELLTAYYNQLCFLDARFLTPAR 187
 241 SLGLFPHWYDLSLTGVPAAQORALAFEGSVLENTGALHTQIGARQDRSCTEGARRAMEAFQ 300
 188 SLGLFPHWYDLSLTGVPAAQORALAFEGSVLENTGALHTQIGARQDRSCTEGARRAMEAFQ 247
 301 RAAGAFSLLENFNSHAPSDPMSAASICALQMLMAQAQCEVFEGLSPASMAPQDCLAQL 360
 248 RAAGAFSLLENFNSHAPSDPMSAASICALQMLMAQAQCEVFEGLSPASMAPQDCLAQL 307
 361 RLAAQAAQAAAYRLVHRTMAQPPVHDYVPSWTLVHVKAIFYRSLAHYHVALCDGS 420
 308 RLAAQAAQAAAYRLVHRTMAQPPVHDYVPSWTLVHVKAIFYRSLAHYHVALCDGS 367
 421 PATEGELPHEQVLEQPPPTSSKPRGVPVLPQELFEERQLGKAHLKRALIQEALRIHALC 480
 368 PATEGELPHEQVLEQPPPTSSKPRGVPVLPQELFEERQLGKAHLKRALIQEALRIHALC 427
 481 RYLREVDLLRAVISQTLQSLAKYAELEDREDDFCAAAEAPDIQPKTHQKPEARMPLRSG 540
 428 RYLREVDLLRAVISQTLQSLAKYAELEDREDDFCAAAEAPDIQPKTHQKPEARMPLRSG 487
 541 KGPDIHRLGPLSVFSAKNRWLVGPVHLTRGEGGFLTLRGDSPVLIATAVPGSQAATA 600
 488 KGPDIHRLGPLSVFSAKNRWLVGPVHLTRGEGGFLTLRGDSPVLIATAVPGSQAATA 547

QY 601 GLKEGDIYIVSNGQPCRWWRHAEVVTTELKAAGEAGASLQVVSLLPSSRLPSLGDPRPVL 660
 DB 548 GLKEGDIYIVSNGQPCRWWRHAEVVTTELKAAGEAGASLQVVSLLPSSRLPSLGDPRPVL 607
 QY 661 GPRGLLRSOREHGCKTPASTWASPRLLNWSRKAQOQKGTGGCPOPCAPVVPAPSSLKHP 720
 DB 608 GPRGLLRSOREHGCKTPASTWASPRLLNWSRKAQOQKGTGGCPOPCAPVVPAPSSLKHP 667
 QY 721 GWP 723
 DB 668 GWP 670
 RESULT 3
 ID AAO18603
 AC AAO18603 standard; protein; 634 AA.
 XX AAO18603;
 XX 24-OCT-2002 (first entry)
 XX Human protein kinase N family kinase related murine protein.
 XX Human; protein kinase N; kinase; PKN; eye retinoblastoma; bocio tumour;
 XX placenta choriocarcinoma; germ cell; leukocyte; uterus tumour; cancer;
 XX pre-B cell acute lymphoblastic leukaemia; Wilm's tumour; enzyme;
 XX brain anaplastic oligodendroma; uterus endometrial adenocarcinoma;
 XX cytostatic; gene therapy; chromosome 8; mouse.
 XX Mus musculus.
 XX WO200261062-A2.
 XX 08-AUG-2002.
 XX 29-JAN-2002; 2002WO-US002152.
 XX 01-FEB-2001; 2001US-00773371.
 XX 07-MAY-2001; 2001US-00849334.
 XX (PEKE) PE CORP NY.
 XX Rusch D, Ketchum KA, Di Francesco V, Beasley EM;
 XX WPI; 2002-608516/65.
 XX New human kinase peptide and nucleic acid molecule, useful for treating
 XX disorders associated with abnormal expression of kinase protein, e.g.
 XX retinoblastoma, Wilm's tumor, in drug screening assays and
 XX pharmacogenomic analysis.
 XX Disclosure; Page 75-76; 76pp; English.
 XX The present invention provides the protein and coding sequences of a
 XX novel human protein kinase N family protein. The sequences can be used in
 XX the treatment of disorders associated with the absence of, inappropriate,
 XX or unwanted expression of the protein, e.g. eye retinoblastoma, Wilm's
 XX tumour, placenta choriocarcinoma, bocio tumours, pre-B cell acute
 XX lymphoblastic leukaemia, uterus tumours, brain anaplastic oligodendromas
 XX and uterus endometrial adenocarcinomas. The present sequence is a murine
 XX protein. Note: The present sequence is not described further in the
 XX exemplification of the invention
 XX Sequence 634 AA;

Query Match 61.8%; Score 2331; DB 5; Length 634;
 Best Local Similarity 67.0%; Pred. No. 6.8e-213;
 Matches 473; Conservative 49; Mismatches 112; Indels 72; Gaps 6;
 QY 1 MILERPPGAGAGESPRLOIQRKPRKTRVSSLRGRREGLRDVCAMRCVRHGEDPVR 60
 DB 1 MILERPPGAGAGESPRLOI-----GSTR----- 26

```

QY 61 VHVGMNPNQLHAVGCDLSLTQCCQLOSRRAQIHQIDKELQMTGAEENLYRATSNRVR 120
DB 27 -----KGYGSFVNQPCQLQSHRARLHQIISKELMRGTGAENLYRATSNWTWR 74

QY 121 ETVALELSYVNSNLQLLKEELEELSGVDPGRHSGEAVTVPMIPLGLKETKELDWSPLK 180
DB 75 ETVALELSYVNSNLQLLKEELAELSTSDVDQPEGEGITIPMIPLGLKETKELDWTPLK 134

QY 181 ELISVHFGEDGASVEAFIRELEALRQAMRTSRNESGLELLTAYYNQLCFDLARFLTPAR 240
DB 135 ELISVHFGEDGTSFETIQELEDLRQATRTPSRDEAGLDLAAIYQSCLFIDARFFSPSR 194

QY 241 SLGLFFHWYDSLTGVPAAQRALAFKGSVLNFIHALHTQIGARQDRSCTEGARAMEAFQ 300
DB 195 SPGLLFFHWYDSLTGVPAAQRALAFKGSVLNFIHALHTQIGARQDRSCTEGTNHAEEAFQ 254

QY 301 RAAGAFSLRLNFSHAPSFDMSAASLCALBQIMMAQAQECVFEGLSPASMAPQDCLAOL 360
DB 255 RAAGAFSLRLNFSHAPSFDMSAASLSMLEQLMIAQAQECIFKGLLLPASATPDICPDQL 314

QY 361 RLAEQAAQVAAYEYRLVHRTMAQPPVHDYVPVSWTALVHVKAIEYFRSLAHYHVAALCDGS 420
DB 315 QLAQAAQVAATEYGLVHRAQAPPPVRDYLPAWNTLHVKAIEHFCALAHYHVAALCESH 374

QY 421 PATEGELETHQVFLQPTSSKPRGPVLPQLEERLQGLKALHRAILQGEALRLHALC 480
DB 375 PA-KGELARQEHVF-QPSTPHEPIGTLTPQHPEDRRKLAKALHRAILQGEALRLHALC 432

QY 481 RVLREVDLLRAVISQTLORSIAKYAELDRDDPCFAAEAPDIQPKTHQKPEARMPRLSQG 540
DB 433 RVLKVDLLQVVVVTQALRRSLAKYSQLEERDDFFATEAPDIQPKTHQTPF 483

QY 541 KGPDIFFHRLGPLSVFSAKNRWELVGPVHLTRGEGGFLTRGDSVPLIAAVIPGSAQAAA 600
DB 484 -----GPLSVFSTKNRWQLVGPVHTRGEGGFLTRGDSVPLIAAVIPGSAQESA 534

QY 601 GLKEGDIYVSVNGQPCRWRRHAEVVTTELKAAAGEAGASLQVSWLLPSSRLPSLGDPRPVL 660
DB 535 GLKEGDIYVSVNGQPCRWKHEVVTQLRSMGEEGVSLQVSWLLPSSPEPRGTGPRRAALL 594

QY 661 GPRGLLRQREHGCKTTPASTWASRPPLNWSRKAQCKTGCGPQRC 706
DB 595 W-----NQREGCFETFPTRTPRPILGWSRKNQCKTGSHDPDC 634

RESULT 4
ABM32685
ID ABM32685 standard; protein; 643 AA.
XX AC ABM32685;
XX DE 06-OCT-2003 (first entry)
XX DT Human Rhophilin amino acid sequence SEQ ID NO:27.
XX DE DGI-3; identification; cytosolic; gene therapy; cancer.
XX KW Homo sapiens.
XX OS WO2003033675-A2.
XX PN 24-APR-2003.
XX PF 18-OCT-2002; 2002WO-US033511.
XX PR 19-OCT-2001; 2001US-0342892P.
XX (DGI-3) DGI BIOTECHNOLOGIES INC.
XX PA Pillutla R, Blume AJ, Goldstein NI;
XX PI WPI; 2003-393525/37.
XX DR
XX

```

```

PT New DGI-3 nucleic acid useful for preparing a composition for diagnosing
or treating cancer.
XX Example 6; Fig 4B; 215pp; English.
PS
CC The present invention describes an isolated nucleic acid (I) comprising:
CC (a) a sequence encoding A1, G3, D9, C5, D1, H2, H3, G12, E7 or A1 or
CC penetrating variants; (b) a sequence comprising at least 15 contiguous
CC nucleotides of (a); or (c) a sequence that is complementary or at least
CC 80% identical to (a) or (b). Also described: (1) a vector comprising the
CC nucleic acid; (2) a virus comprising the nucleic acid; (3) a host cell
CC comprising the vector; (4) a primer or probe comprising the nucleic acid;
CC (5) an isolated peptide comprising the amino acid sequence of A1, G3, D9,
CC C5, D1, H2, H3, G12, E7 or A1 or penetrating variants; (6) an antibody
CC that binds to the isolated peptide; (7) an isolated polypeptide complex
CC comprising the peptide and DGI-polypeptide complex; (8) a peptide library
CC comprising the peptide; (9) identifying a DGI-3 binder; (10) identifying
CC a DGI-3 antagonist agent; (11) diagnosing cancer; (12) a pharmaceutical
CC composition; (13) treating cancer; and (14) inhibiting activity in a
CC mammalian cell of a gene whose activity is regulated by DGI-3. (i) has
CC cytosolic activity and can be used in gene therapy. The nucleic acid (I)
CC can be used for preparing a composition for diagnosing or treating
CC cancer. ACF56944 to ACF57031 and ABM32656 to ABM32775 represent sequences
CC used in the exemplification of the present invention
XX
SQ Sequence 643 AA;

```

```

Query Match 61.8%; Score 2331; DB 6; Length 643;
Best Local Similarity 67.0%; Pred. No. 7e-213;
Matches 473; Conservative 49; Mismatches 112; Indels 72; Gaps 6;

QY 1 MILERPPGACAGBESPRLOISRRKPRKTRVSSURRREGRLDVCAMGRCVRHGEDPVR 60
DB 1 MILERPPGQGTGESSRPQDD-----GSTR----- 26

QY 61 VHVGMNPNQLHAVGCDLSLTQCCQLOSRRAQIHQIDKELQMTGAEENLYRATSNRVR 120
DB 27 -----KGYGSFVNQPCQLQSHRARLHQIISKELMRGTGAENLYRATSNWTWR 74

QY 121 ETVALELSYVNSNLQLLKEELEELSGVDPGRHSGEAVTVPMIPLGLKETKELDWSPLK 180
DB 75 ETVALELSYVNSNLQLLKEELAELSTSDVDQPEGEGITIPMIPLGLKETKELDWTPLK 134

QY 181 ELISVHFGEDGASVEAFIRELEALRQAMRTSRNESGLELLTAYYNQLCFDLARFLTPAR 240
DB 135 ELISVHFGEDGTSFETIQELEDLRQATRTPSRDEAGLDLAAIYQSCLFIDARFFSPSR 194

QY 241 SLGLFFHWYDSLTGVPAAQRALAFKGSVLNFIHALHTQIGARQDRSCTEGARAMEAFQ 300
DB 195 SPGLLFFHWYDSLTGVPAAQRALAFKGSVLNFIHALHTQIGARQDRSCTEGTNHAEEAFQ 254

QY 301 RAAGAFSLRLNFSHAPSFDMSAASLCALBQIMMAQAQECVFEGLSPASMAPQDCLAOL 360
DB 255 RAAGAFSLRLNFSHAPSFDMSAASLSMLEQLMIAQAQECIFKGLLLPASATPDICPDQL 314

QY 361 RLAEQAAQVAAYEYRLVHRTMAQPPVHDYVPVSWTALVHVKAIEYFRSLAHYHVAALCDGS 420
DB 315 QLAQAAQVAATEYGLVHRAQAPPPVRDYLPAWNTLHVKAIEHFCALAHYHVAALCESH 374

QY 421 PATEGELETHQVFLQPTSSKPRGPVLPQLEERLQGLKALHRAILQGEALRLHALC 480
DB 375 PA-KGELARQEHVF-QPSTPHEPIGTLTPQHPEDRRKLAKALHRAILQGEALRLHALC 432

QY 481 RVLREVDLLRAVISQTLORSIAKYAELDRDDPCFAAEAPDIQPKTHQKPEARMPRLSQG 540
DB 433 RVLKVDLLQVVVVTQALRRSLAKYSQLEERDDFFATEAPDIQPKTHQTPF 483

QY 541 KGPDIFFHRLGPLSVFSAKNRWELVGPVHLTRGEGGFLTRGDSVPLIAAVIPGSAQAAA 600
DB 484 -----GPLSVFSTKNRWQLVGPVHTRGEGGFLTRGDSVPLIAAVIPGSAQESA 534

QY 601 GLKEGDIYVSVNGQPCRWRRHAEVVTTELKAAAGEAGASLQVSWLLPSSRLPSLGDPRPVL 660
DB 535 GLKEGDIYVSVNGQPCRWKHEVVTQLRSMGEEGVSLQVSWLLPSSPEPRGTGPRRAALL 594

```

RESULT 6	
AAB43571	
ID	AAB43571 standard; protein; 257 AA.
XX	
XX	AAB43571;
XX	
XX	08-FEB-2001 (first entry)
XX	
XX	Human cancer associated protein sequence SEQ ID NO:1016.
DE	
XX	
XX	Human; cancer associated gene; cancer antigen; detection; cancer;
XX	diagnosis; cystostatic; proliferative; vulnery; immunomodulator;
KW	antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW	antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
KW	dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW	vasotrophic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW	immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW	allergic reaction; graft versus host disease; organ rejection;
KW	haemostatic; thrombolytic; cardiovascular disorder; infection;
KW	neurological disease; drug screening.
XX	
OS	Homo sapiens.
XX	
XX	WO200055350-A1.
PN	
XX	
XX	21-SEP-2000.
PD	
XX	
XX	08-MAR-2000. 2000WO-05005982
PF	

```

XX 12-MAR-1999; 99US-0124270P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM;
XX WPI; 2000-587533/55.
XX N-PSDB; AAC77780.
XX Novel isolated nucleic acids comprising sequences encoding peptides
XX useful for treating or diagnosing e.g. cancer.
XX Claim 11; Page 1597-1598; 2352pp; English.
XX
XX AAC77607 to AAC78448 encode the human cancer associated proteins given in
XX AAB43398 to AAB44239. The proteins can have activities based on the
XX tissues and cells the genes are expressed in. Example of activities
XX include: cytostatic; proliferative; vulnerary; immunomodulator;
XX antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
XX antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
XX dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
XX nototropic; vasotropic; antipsoriatic and antiangiogenic. The
XX polynucleotides and polypeptides can be used for preventing, treating or
XX ameliorating medical conditions and diagnosing pathological conditions.
XX Polynucleotides, polypeptides, antibodies, agonists and antagonists from
XX the present invention may be used to treat immune disorders by activating
XX or inhibiting the proliferation, differentiation or mobilisation of
XX immune cells; to treat disorders of haematopoietic cells, autoimmune
XX disorders, allergic reactions, graft versus host disease and organ
XX rejection, modulate haemostatic or thrombolytic activity, modulate
XX inflammation, cancers, cardiovascular disorders, neurological disease and
XX bacterial or viral infections. The peptides, nucleotides, antibodies,
XX agonists and antagonists may be also be used in drug screens. AAC78449 to
XX AAC78457 and AAB4240 represent sequences used in the exemplification of
XX the present invention
XX
XX Sequence 257 AA;
XX
XX Query Match 33.9%; Score 1278; DB 3; Length 257;
XX Best Local Similarity 99.2%; Pred. No. 5.9e-113;
XX Matches 245; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 459 GKAKHLKRAILGQEEALRHLCRVLRVDLLRAVISQTLQRSIAKYAELDRDDFCEAAE 518
Db 8 GKAKHLKRAILGQEEALRHLCRVLRVDLLRAVISQTLQRSIAKYAELDRDDFCEAAE 67
QY 519 APDIQPKTHQKPEARMPRLSQKGPDI FHLRGLPLSVFSKAKNRWLVGPVHLTRGEGGFL 578
Db 68 APDIQPKTHQKPEARMPRLSQKGPDI FHLRGLPLSVFSKAKNRWLVGPVHLTRGEGGFL 127
QY 579 TLRGDSPVLIAAVIPGSAAGLKEGDYIVSVNGQPCRWHRHAEVVTLKAAGAGASL 638
Db 128 TLRGDSPVLIAAVIPGSAAGLKEGDYIVSVNGQPCRWHRHAEVVTLKAAGAGASL 187
QY 639 QVVSLLPSRLPSLGRPVLLGPRLLRSQREHGCKTPASTWASPRLLNWSRKAQQK 698
Db 188 QVVSLLPSRLPSLGRPVLLGPRLLRSQREHGCKTPASTWASPRLLNWSRKAQQK 247
QY 699 TGGCQFP 705
Db 248 TGGCQFP 254
XX
RESULT 7
AAU00869
ID AAU00869 standard; protein; 257 AA.
XX
XX AAU00869;
XX AC
XX XX
XX 04-JUL-2001 (first entry)
XX
XX Human cancer related protein 5.

```

```

XX Human; cancer related protein; food additive; preservative; immunogen;
KW antibody; bone cancer; adrenal cancer; bone marrow cancer; breast cancer;
KW gastrointestinal cancer; liver cancer; lung cancer; urogenital cancer;
KW immune disorder; Addison's disease; allergy; diabetes mellitus;
KW autoimmune haemolytic anaemia; autoimmune thyroiditis; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW acquired immunodeficiency syndrome; AIDS; cardiovascular disorder;
KW myocardial ischaemia; wound healing; neurological disorder;
KW Parkinson's disease; Alzheimer's disease; cerebral anoxia; epilepsy;
KW viral infection; bacterial infection; fungal infection;
KW parasitic infection; agonist; antagonist.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Region 69..81
XX /label= Immunogenic_epitope
XX Region 84..91
XX /label= Immunogenic_epitope
XX Region 106..112
XX /label= Immunogenic_epitope
XX Region 216..224
XX /label= Immunogenic_epitope
XX Region 239..250
XX /label= Immunogenic_epitope
XX
XX WO200118014-A1.
XX 15-MAR-2001.
XX
XX 30-AUG-2000; 2000WO-US23794.
XX
XX 03-SEP-1999; 99US-0152296P.
XX 06-OCT-1999; 99US-0158003P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Roschke V;
XX
XX WPI; 2001-235186/24.
XX N-PSDB; AAS00831.
XX
XX Twenty nine nucleic acid molecules encoding human cancer associated
XX proteins, useful in the prevention, treatment and diagnosis of cancer,
XX immune disorders, cardiovascular disorders and neurological diseases.
XX
XX Disclosure; Page 394-395; 427pp; English.
XX
XX The sequence represents a novel Human cancer related protein. The
XX polynucleotides and polypeptides are useful for preventing, treating or
XX ameliorating a medical condition in e.g. humans, mice, rabbits, goats,
XX horses, cats, dogs, chickens or sheep. The polypeptides can also be used
XX as a food additive or preservative to increase or decrease storage
XX capabilities. The polynucleotide are useful for chromosome
XX identification. The nucleic acids, protein, antibodies, agonists and
XX antagonists are useful in the diagnosis, treatment and prevention of
XX cancer (e.g. cancers of the adrenal gland, bone, bone marrow, breast,
XX gastrointestinal tract, liver, lung, or urogenital), immune disorders
XX (e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,
XX autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
XX sclerosis, rheumatoid arthritis and ulcerative colitis, acquired
XX immunodeficiency syndrome, AIDS), cardiovascular disorders such as
XX myocardial ischaemias, wound healing, neurological diseases (e.g.
XX Parkinson's disease, Alzheimer's disease, cerebral anoxia and epilepsy)
XX and infectious diseases such as viral, bacterial, fungal and parasitic
XX infections. Numerous examples of each type of disorder are given in the
XX specification
XX
XX Sequence 257 AA;
XX
XX Query Match 33.9%; Score 1278; DB 4; Length 257;
XX Best Local Similarity 99.2%; Pred. No. 5.9e-113;

```


Db 24 GCNPLAQTGRSKLQNRALNQIILKAVRMRTGAENLLKVAATNSKVRQVRLSLFVNSD 83
QY 134 LQLKKEELEELSGGVDPRGHGSEAVTVPMIPLGLKETKELDWSTPLKELISVHFGECDGAS 193
Db 84 LQMLKEELEGLNISGVYQNTTEAFTIPLIPLGLKETKDVDFAVVVKDFILEHSEGYL 143
QY 194 YEAEIRELEALFOAMRTSRNESGLELITAYYNQLCFDARFLTPARSIGLFFHWDYSLT 253
Db 144 YEDEIADLMDLRQACRTSRDEAGVELLMTYFIQLGFVESRFFPTTRQMGLLFTWYDSLT 203
QY 254 GVPAAQALAFKGSVLNIGALHTQIGARODRSCTEGARRAMEAFQAAAGAFSLLENF 313
Db 204 GVPVQQNLLLEKASVLENTGALYTIQIGTRCDRQTOAGLESADAFQAAAGVNLKDTF 263
QY 314 SHAPSDMSAASLCALBOLMAQAQECVFEGLSPPASMAPODCLAQLRLAQAQAAVAAEY 373
Db 264 THTPSYDMSAPMLSVLVKMLAQAQESVFEXISLPG--IRNEFFMLVKVAQAQAAVGEVY 321
QY 374 RLVHRTMAQPPVHDYVPVSWTALVHVKAEPYFRSLAHYHVMALCDG--SPATEGELPTHE 431
Db 322 QQLHAAMSQAPVKENIPYSWASLACVKAHHVAALAHYFTAILIDHQVKPGTDLD--HQ 378
QY 432 QVFLQPTSSKPRG--PVLPP-QLERERQLGKAHLKRAILGOEALRLHALCRVLEVDL 488
Db 378 EKLSQLYDHMPGEGTLPLATLKNQDQRRQLGKSHLRRAAMAHHEESVREASLCKLRSIEV 438
QY 489 LRAVISQTLQSLAKYAEILDREDDFCEAAEAPDIQPKTHQKPEARMPRLSQCKGPDIFHR 548
Db 439 LQKVLCAQAQERSRLTYAQHQBEDDLNLIDAPSVVAKTEQEVDIILPQFSKLTWTDFFQK 498
QY 549 LGPLSVFSKAKNRWRLVGPVHLTRGGFGGLTLRGDSPVLIAAVIPGSAQAAGLKEGDIY 608
Db 499 LGPLSVFSANKRWTPPISRFTAEEDGLGFTLRGNAPVQVHFLDPYCSAVAGAREGDIY 558
QY 609 VSVNGQPCRWHRHAFVVTTELKAAAGASLQVVSLLPSS 647
Db 559 VSIQLVDCWKLTLSVMKMLKXSFGEDETEMKVVSLLDST 597

RESULT 9

AAU10192

ID AAU10192 standard; protein; 686 AA.

XX AC

AAU10192;

XX DT

16-JAN-2002 (first entry)

XX DE

Human prostate specific protein PSL22.

XX KW

Human; prostate specific protein; PSL22; prostate cancer;
benign prostatic hyperplasia; acute prostatitis; testicular cancer;
cryptorchidism; testicular disorder; proliferative disorder; lymphoma;
leukemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;
liver cancer; lung cancer; cytostatic.

XX OS

Homo sapiens.

XX PN

WO200172962-A2.

XX PD

04-OCT-2001.

XX PF

23-MAR-2001; 2001WO-US009410.

XX PR

24-MAR-2000; 2000US-0191929P.

XX PA

(SAAT/) SAATCIOGLU F.

XX PI

Saatecioglu F;

XX DR

WPI; 2001-662926/76.

XX DR

N-PSDB; AAS15827, AAS15828.

XX PT

New polynucleotide for the diagnosis, prevention and treatment for

PT prostate and testis disorders, particularly prostate cancer, comprises
prostate-specific or testis-specific nucleic acids.

XX Claim 1; Fig 14B; 114pp; English.

XX The invention relates to substantially pure prostate-specific or testis-
specific polypeptides and the nucleic acids encoding them. Also included
are vectors and host cells expressing the proteins, a transgenic animal
CC expressing the protein, antibodies against the proteins, probes for
CC detecting the nucleic acids, antisense molecules for the nucleic acids
CC and methods of isolating modulators of the proteins. Compounds that
CC modulate the prostate specific or testis specific polypeptide are useful
CC to diagnose, prevent or treat disorders of the testis or prostate
CC particularly prostate cancer, benign prostatic hyperplasia, acute
CC prostatitis, testicular cancer, cryptorchidism, undescended, retractile,
CC ascending or vanished testis. Other proliferative disorders for which the
CC modulators may be used include lymphoma, leukaemia, melanoma, ovarian
CC cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer.
XX The present sequence represents a prostate specific protein, PSL22

SQ Sequence 686 AA;

Query Match 33.3%; Score 1255; DB 4; Length 686;

Best Local Similarity 45.9%; Pred. No. 4.6e-110;

Matches 266; Conservative 107; Mismatches 196; Indels 10; Gaps 5;

QY 74 GCDLSLTQICQGLQSRRAQIHQQIDKELQMRGTGAENLYRATNSNRVRETVALELSYVNSN 133

Db 24 GCNPLAQTGRSKLQNRALNQIILKAVRMRTGAENLLKVAATNSKVRQVRLSLFVNSD 83

QY 134 LQLKKEELEELSGGVDPRGHGSEAVTVPMIPLGLKETKELDWSTPLKELISVHFGECDGAS 193

Db 84 LQMLKEELEGLNISGVYQNTTEAFTIPLIPLGLKETKDVDFAVVVKDFILEHSEGYL 143

QY 194 YEAEIRELEALFOAMRTSRNESGLELITAYYNQLCFDARFLTPARSIGLFFHWDYSLT 253

Db 144 YEDEIADLMDLRQACRTSRDEAGVELLMTYFIQLGFVESRFFPTTRQMGLLFTWYDSLT 203

QY 254 GVPAAQALAFKGSVLNIGALHTQIGARODRSCTEGARRAMEAFQAAAGAFSLLENF 313

Db 204 GVPVQQNLLLEKASVLENTGALYTIQIGTRCDRQTOAGLESADAFQAAAGVNLKDTF 263

QY 314 SHAPSDMSAASLCALBOLMAQAQECVFEGLSPPASMAPODCLAQLRLAQAQAAVAAEY 373

Db 264 THTPSYDMSAPMLSVLVKMLAQAQESVFEXISLPG--IRNEFFMLVKVAQAQAAVGEVY 321

QY 374 RLVHRTMAQPPVHDYVPVSWTALVHVKAEPYFRSLAHYHVMALCDG--SPATEGELPTHE 431

Db 322 QQLHAAMSQAPVKENIPYSWASLACVKAHHVAALAHYFTAILIDHQVKPGTDLD--HQ 378

QY 432 QVFLQPTSSKPRG--PVLPP-QLERERQLGKAHLKRAILGOEALRLHALCRVLEVDL 488

Db 378 EKLSQLYDHMPGEGTLPLATLKNQDQRRQLGKSHLRRAAMAHHEESVREASLCKLRSIEV 438

QY 489 LRAVISQTLQSLAKYAEILDREDDFCEAAEAPDIQPKTHQKPEARMPRLSQCKGPDIFHR 548

Db 439 LQKVLCAQAQERSRLTYAQHQBEDDLNLIDAPSVVAKTEQEVDIILPQFSKLTWTDFFQK 498

QY 549 LGPLSVFSKAKNRWRLVGPVHLTRGGFGGLTLRGDSPVLIAAVIPGSAQAAGLKEGDIY 608

Db 499 LGPLSVFSANKRWTPPISRFTAEEDGLGFTLRGNAPVQVHFLDPYCSAVAGAREGDIY 558

QY 609 VSVNGQPCRWHRHAFVVTTELKAAAGASLQVVSLLPSS 647

Db 559 VSIQLVDCWKLTLSVMKMLKXSFGEDETEMKVVSLLDST 597

RESULT 10

ABG94709

ID ABG94709 standard; protein; 686 AA.

XX AC

ABG94709;

XX XX

DT 02-DEC-2002 (first entry)
XX Human GTP-Rho binding protein 2.
XX
KW Human; GTP-Rho binding protein 2; GRBP2; chromosome 19q12; antigen;
KW oncogene; tumour; liposarcoma; ichthyosis congenita III; vaccine;
KW benign familial infantile convulsion; gene therapy.
XX
OS Homo sapiens.
XX
XX EP1231216-A2.
XX
XX 14-AUG-2002.
XX
XX 17-JAN-2002; 2002EP-00001026.
XX
XX 30-JAN-2001; 2001WO-US000663.
XX 30-JAN-2001; 2001WO-US000664.
XX 30-JAN-2001; 2001WO-US000665.
XX 30-JAN-2001; 2001WO-US000666.
XX 30-JAN-2001; 2001WO-US000667.
XX 30-JAN-2001; 2001WO-US000668.
XX 30-JAN-2001; 2001WO-US000669.
XX 30-JAN-2001; 2001WO-US000670.
XX 29-JUN-2001; 2001US-00895040.
XX (AEOM-) AEOMICA INC.
XX
XX Shannon ME, JI Y;
XX WPI; 2002-684026/74.
XX P-PSDB; ABS71855, ABS71856.
XX
XX Novel GTP-Rho binding protein 2 and nucleic acids encoding the protein,
XX useful for the manufacture of a medicament for treating a disease
XX associated with altered expression or activity of human GRBP2 protein.
XX
XX Claim 10; Fig 3; 101pp; English.
XX
XX The invention relates to an isolated GTP-Rho binding protein 2 (GRBP2)
XX polypeptide or a fragment of at least 6 amino acids or a sequence in
XX which at least 95% of deviations from GRBP2 sequences are conservative
XX substitutions. Also included are an isolated nucleic acid (GRBP2 NA)
XX encoding GRBP2 comprising the full length cDNA or CDS, fragments or
XX variants, GRBP2 vectors, host cells, antibodies, transgenic non-human
XX animals modified to contain GRBP2 NA (or unable to express the endogenous
XX orthologue of GRBP2), diagnosing a disease caused by a mutation in human
XX GRBP2 or altered expression of GRBP2, ant-agonists of GRBP2, GRBP2
XX microarrays, fusion proteins and screening for agents that modulate the
XX expression of GRBP2 NA. GRBP2 is useful for identifying binding partners
XX of GRBP2. GRBP2 NA and Ab are useful in therapy and in the
XX manufacture of a medicament for the treatment or prevention of a disorder
XX associated with increased or decreased expression or activity of human
XX GRBP2 (e.g. tumours, liposarcoma, ichthyosis congenita III and benign
XX familial infantile convulsion, all associated with the chromosomal
XX location of GRBP2, 19q12). GRBP2 is useful as a standard in immunoassay
XX specific for the proteins, to be used in a therapeutic agent, as
XX vaccines, to be and as antigens (e.g. for epitope mapping) or immunogens
XX (e.g. for raising antibodies). GRBP2 NA is useful as hybridisation probes,
XX to prime synthesis of nucleic acids, to prime first strand cDNA sequence,
XX on an mRNA template, and to drive in vivo expression of the proteins. The
XX vector is useful for shuttling GRBP2 NA between host cells derived from
XX disparate organisms, for inserting GRBP2 NA into host cell chromosome,
XX for expressing sense or antisense RNA transcripts of GRBP2 NA in vitro or
XX within a host cell, and for expressing GRBP2 alone or as fusions to
XX heterologous polypeptides. The antibody is useful as an analytical
XX reagent for detection and quantification of GRBP2 and as an immuno
XX therapeutic agent and is useful for flow cytometric detection, for
XX scanning laser cytometric detection, or for fluorescent immunoassay. The
XX present sequence is a GRBP2 protein sequence
XX
XX Sequence 686 AA;

Query Match 33.1%; Score 1249; DB 5; Length 686;
Best Local Similarity 45.8%; Pred. No. 1-7e-109;
Matches 265; Conservative 107; Mismatches 197; Indels 10; Gaps 5;

QY 74 GDSLTQIOCGLOSPRAHQHIOIDKELQMRCAENLYRATSNRRVRETVALELSVNSN 133
DB 24 GGNPLAQTGRSLQNRALNQOILKAVRMIRGAENLLKVATNSKVRREQRLBELSFVNSD 83
QY 134 LQLKEELEELSGGVDPRGRHGEAVTPMIPICLKETKELDMSTPLKELISVHFGEDEGAS 193
DB 84 LQMLKEELEGLNISVGVIQNTTEAFTIPLIPGLKETKDVDFAVVLKDFILEHYSDEGYL 143
QY 194 YEABIREALRQAMRTPSRNSGLELLTAYYNQLCFDARFLTPARSLGLFFHMYDSLT 253
DB 144 YEDEIADMLDRQACRTPSRDEAGVELLMTYFQLGFVESRFPPTQMGILLFTWYDSLT 203
QY 254 GYPAQORALAFKESVLENIQALHTOIGARQDRSCTEGARRAMEAFORAGAFSLIRENF 313
DB 204 GYFVSQNNLLLEKASVLENTGALYIQIGTRCDRQTQAGLESALDAFORAAGVINYLKDTF 263
QY 314 SHAPSDMSAASICALQELMMAQAQCEVFEGLSPFPASMAPQDCLAQLRLAQEAQAAVEY 373
DB 264 THTPSYDMSPAVLVILVKMLAQAESVFEKISLPG--IRNEFFMLVKYAEAAKGEVY 321
QY 374 RLVRHRTMAQPPVHDVPSWTALVHVKAIEYFRSLAHYHVAMALCDG--SPATEGELPTHE 431
DB 322 QQLHAAMSQAQPKENITPYSWASLACVKAHYAALAHYFTAILIIDHQVKPGTDLD---HQ 378
QY 432 QVFLQPTTSSKPRG--PVLP-QLERROLGKAHLKRAILGQEEALRLHALCVLREVDL 488
DB 379 EXCLSQLYDHMPGLTFLATLKNQQRQLGSHLURAWAHHEESVREASLCKKLSIEV 438
QY 489 LRRAVISQTLQRSIAKAEALDREDDFCEAAEAPDIOPKTHQKPEARMPLRSQKGGPDIHFR 548
DB 439 LQKVLCAAQERSRLTYAQHQEEDDLNLIDASVVAKTEQEVDDIILPQFSKLTVTDFQK 498
QY 549 LGPLSVFSKXNRWLVGPVHLTRGEGGFGLTIRGDSFVLIAAVIPGSAQAAGLKEGDYI 608
DB 499 LGPLSVFSANKWTTPRSIRFTAEEGDLGFTLRGNAPQVHFLLDPYCSASVAGAREGYI 558
QY 609 VSVGQPCRWHRHAEEVTVTELKAAGSAGASLOVVSLPSS 647
DB 559 VSIQLVDCXKWLTLSEVMKLLKSFGEDETEMKVSLDST 597

RESULT 11
AAB68561
ID AAB68561 standard; protein; 666 AA.
AC AAB68561;
XX
DT 19-APR-2001 (first entry)
XX Human GTP-binding associated protein #61.
DE
XX
KW Human; guanosine triphosphate binding associated protein; GTP; GRAP;
KW inflammation; AIDS; Addison's disease; anaemia; arteriosclerosis; asthma;
KW autoimmune disorder; hepatitis; multiple sclerosis; cancer; diabetes;
XX osteoporosis; psoriasis.
OS Homo sapiens.
XX
XX WO200105970-A2.
XX
XX 25-JAN-2001.
XX
XX 19-JUL-2000; 2000WO-US019698.
XX
XX 19-JUL-1999; 99US-0144595P.
XX 23-AUG-1999; 99US-0150460P.
XX 15-OCT-1999; 99US-0159849P.
XX
XX (INCY-) INCYTE GENOMICS INC.

XX Yue H. Tang YT, Bandman O, Hillman JL, Lal P, Au-Young J;
PI Reddy R, Yang J, Baughn MR, Lu DAM, Azimzai Y, Patterson C;
XX WPI; 2001-091972/10.
DR N-ESDB; AAF58361.
XX New guanidine triphosphate-binding associated proteins (GBAP) and their
PT encoding nucleic acids, useful for treating and/or diagnosing diseases
PT associated with GBAP expression, such as cancer, diabetes and asthma.
XX Claim 1; Page 180-181; 233pp; English.
PS The present invention relates to novel human guanidine triphosphate (GTP)
CC -binding associated proteins (GBAPs; AAB68501-AAB68566) and their coding
CC sequences (AAF58301-AAF58366). The proteins and coding sequences of the
CC present invention are useful for treating a variety of disorders
CC including inflammation, AIDS, Addison's disease, anaemia,
CC arteriosclerosis, asthma, autoimmune disorders, Grave's disease,
CC hepatitis, multiple sclerosis, cancer, diabetes, osteoporosis and
XX psoriasis
SQ Sequence 666 AA;
Query Match 32.8%; Score 1236; DB 4; Length 666;
Best Local Similarity 46.1%; Pred. No. 2.9e-108;
Matches 262; Conservative 106; Mismatches 190; Indels 10; Gaps 5;
QY 85 QLOSRAHQIDKELQMRGTGAENLYATSNRRVREVALELSVNSNLQILKEEL 144
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
15 KIQNQAALNQOILKAVRMRTGAENLLKVATNSKVRQVRLSEFVNSDLQMLKEEL 74
QY 145 SGGVDPGRHGSFAVTVPMLPLGKTKELDNSTPLKELISVHFGEDGASVYAEIRELEAL 204
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
75 NLSGVYQNTAEFTPIPLIGLKKTKVDFAVLKDFILEHSEGVLYEIDEADLMDL 134
QY 205 ROAMETPSRNEGLELLTAYYNQLCFDLARFTTPARSLGFPWHYDSTGVPAQRAIAF 264
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
135 RQACTPSRDEAGVELLTFYIQLGFVSRFPPTTRQMGLLFTWYDSLTVGPVSQNLIL 194
QY 265 EKGSVLFTNIGALHTOIGARQDSCTEGARRAMEAFQRAAGFSLARENFSHAPSMDMAA 324
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
195 EKASVLFNTGALYQIGTRCDROTQAGLESADAFQRAAGVNLKDTFTHTPSYDMSPA 254
QY 325 SICALEQLNMAQAQECVFEGLSPFAPQDCCLAQRLAQEAQAAYAAEYRLVHRTMAQPP 384
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
255 MLSVLVKMLAQASVFEKISLPG--IXNEFFMLVKVAQEAQKGVYQQLHAAMSQAP 312
QY 385 VHDYVPVSWTALVHVAEYFSLAHYHVMALCDG--SPATSEGLPTHEQVFLQPTTSK 442
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
313 VKENIPYSWASLACVKAHYAALAHYFTAILLIDHQVKEGTDLD--HOEKCLISQLYDHM 369
QY 443 PRG--PVLP--QELESRRQIGKALHRAILGQEAELRHLCRLRVLDLRAVISQTLQR 499
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
370 PEGLLPLATLKNQDQQRQKGRKLRMAHHEESVREASLCRKARTIEVLQKVLCAQER 429
QY 500 SLAKVAELDREDDFCEAABAPIQKTHQKPEARMPRLSQGKGPDIHFRLGLPSVFSAXN 559
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
430 SRLTAHQEQEEDLLNLDAPSVAKTEQEVDIILPQFSKLTWTDFFQKGLPSVFSANK 489
QY 560 RWRVLGVPVHLTREGGFGTLRGDSPVLTAAVPGSQAAAAGLKEGDYIVSVNGQPCRW 619
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
490 RWTTPRSIRFTABEGDLGTLRGNAFVQVHFIDPYCSASVAGAREGDYIVSIQLVCKWL 549
QY 620 RHAEVVTELKAAAGASQVVSLLPSS 647
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
550 TLSEVWKLKSGFEDEMKVVSLLDST 577

RESULT 12
ABU65215
ID ABU65215 standard; protein; 686 AA.
XX

AC ABU65215;
XX 20-MAY-2003 (first entry)
DT Human NOV126a protein.
DE NOYX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
XX hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW human.
KW Homo sapiens.
OS WO200272757-A2.
XX 19-SEP-2002.
XX 08-MAR-2002; 2002WO-US006908.
PR 08-MAR-2001; 2001US-0274101P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 08-MAR-2001; 2001US-0274322P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-0276776P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0277338P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.

PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332894P.
PR 03-DEC-2001; 2001US-0337426P.
PR 04-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.
XX (CURA-) CURAGEN CORP.
XX
PI Padigaru M, Spytek Ka, Shenoy SG, Taupier RJ, Pena CEA, Li L;
PI Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangoli E, Vernet CAM, Guo X, Tchernev V;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
PI Lepley DM, Rieger DK;
DR WPI; 2002-723332/78.
DR N-PSDB; ABX97182.
XX
XX
XX
PT NOVX polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVX expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
PT asthma.
XX
XX
PS Claim 1; Page 470-471; 1103pp; English.
XX
XX This invention describes novel human NOVX polypeptides which have
CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive
CC activity. Pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. AB065041-AB065218 represent the NOVX polypeptides encoded by
CC ABX97008-ABX97185
XX
SQ Sequence 686 AA;

Query Match
Best Local Similarity 31.7%; Score 1195; DB 5; Length 686;
Matches 263; Conservative 104; Mismatches 196; Indels 22; Gaps 8;

QY 74 GDSLTQICGQSQSPRAHQIQIDKELQMRGAENLYRATSNRNVRETVALELSVNSN 133
DB 24 GNPQAQTRGSLQNRALNQOILKAVRMRTGAENLLKVATNSKVRQVRLELSPVNSD 83
QY 134 LQLKKEELELGGVDPGRHGEAVTPMPLGLKETKELDSTPLKELISVHFGEKGAS 193
DB 84 LQMLKEELEGLNISGVVQNTTEATIPPLGLKETKQDVFAVLKQFLEHSEGGYL 143
QY 194 YBAETRELEALQAMRTPSNBSGLELLTPAYNQLCFDLARELTTPARSIGLFFHWYDSLT 253
DB 144 YEDEIADLMDLQACRTSRDEAGVELLMTYFIQLGFVESRFPPTQRMGLLFTWYDSLT 203
QY 254 GYPAQORALAFKSGVLENGALHTQIGAROBSCTEGARAMEAPQRAAGFALLRENF 313
DB 204 GVPVQQNLLLEKASVLFNTGALYTOIGTRCDRQTOAGLESIDAIPORAAGVNLVKTDF 263
QY 314 SHAPSPDMSAALCALQELMLMAQAQECVPEGLSPASMAPQCLAQRLQAQAAQVAA 373
DB 264 THTPSYDMSPALSVLVKMLQAQESVEKISLPG--IRNFFMLVKVQAQAAGVEVY 321
QY 374 RLVHRTMAQPPVHDVYVPSWNTALVHVKAEIFRSLAHYHVAMALCDG--SPATEGELP 431
DB 322 QQLHAAMSQAPKENIPYSWASLACVKAHYAALAHYFTAILLDHVKVPGTDLD--HQ 378
QY 432 QVFLQPTSSKPRG--PVLFP-QELBERQLGKHLKRALGQBEALRLHALCRVREVDL 488
DB 379 EKLCSQLYDHPMPEGLTFLATLKNQDQRRQLGKSHLFRAMAHHEESVREASLCKLRSIEV 438
QY 489 LRAVISQTLQRLSLAKVAELDRDDDFCEAAEAAPDI-----QPKTHQKPEARMLSQGKG 542

DB 439 LQKVLCAAQERSRLTYAQHQEEDLNLNLDAPRVLLKLKSLTYCPSS--PADSHGLL 496
QY 543 POIFHRLGPLSVFSKAKNRWLVGPVHLTRGEGGFGTLTRGDSPLVLTAAVIRGSAAGL 602
DB 497 FE----AGPLSVLSANKRWTPFRSIRFTAEBGDLGFTLRGNAFVQVHFLDPYCSAVAGA 552
QY 603 KEGDYIVSYNGQPCRWRRHAEVVTELKAAAGEAGASLQVVSLLPSS 647
DB 553 REGDYIVSIQLVDCKWLTLSEVMKLIKSGFGEDEIEMKVVSLDST 597
RESULT 13
ABG08369
ID ABG08369 standard; protein; 1111 AA.
XX
AC ABG08369;
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #8360.
XX
KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS72556.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 38728; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1111 AA;

RESULT 15

ABB63879
ID ABB63879 standard; protein; 648 AA.

XX AC ABB63879;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 18429.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX KW pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEXE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX DR N-PSDB; ABL07982.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signaling and cell-cell
XX interactions.

XX Disclosure; SEQ ID NO 18429; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 648 AA;

Query Match 21.6%; Score 817; DB 4; Length 648;

Best Local Similarity 31.7%; Pred. NO. 2.9e-88;

Matches 215; Conservative 114; Mismatches 233; Indels 116; Gaps 19;

QY 98 DKELQWRTGAENLYRATSNRRVRETVL-ELSY- ---VNSNLQLLKEELEELSGGVDPGR 152

DB 3 EKEHELST-----RVVKNEENDEKLNELSEFCFVRINSSVEIYQSE----- 44

QY 153 HGSEAVTVPMPILGKTKELDWSTPLKELISVHFGEDGASYEAEIREALRQAMRTPS 212

DB 45 --SHNGIMPMIPLGLKETKEINFMPEPDSFILEHYSEEPSMYIDATADMTDTRQASKTPS 102

QY 213 RNESGLELITAYNQICFLDARFLTPARSIGLFFHWYDSLTGVPQORALAFKGSVLN 272

DB 103 RDALGVALLFRYNTLYYVERFFPPDRNLGVYFEWYDSLTGVPSCQRTIAFEKACTLEN 162

QY 273 IGAHTQIGARODRSTEGARAMEAFORAAGAFSLLENFESHAPSPDMSAASLCALEOL 332

DB 163 LGGIYTIQIGARHRTTRTERGLDIAVDFLRAAGVFRHIYDTFTNPSMDLKPQVLDVNL 222

QY 333 MMAQAQECVEGLSPSPASMAPQDCLAQRLAQEAQAAAEYRLVHRTMAQPPVHDYVPVS 392

DB 223 MLSQARECLFEKLQLQIEAMSHDCQAFRLDAGEAAQISHEYNEMHKNIQANDHTIYLPFC 282

QY 393 WTALVHVKAERYFRSLAHYHVAMALCDGSPATEGELPHEQVFLQPTTSKPRGEVLPQEL 452
DB 283 WAGLVFVKAELYKAPAFYKARS-----DATD-ELKASK-----SSQKNQSSFIGNSQ 330
QY 453 EERR-----QLGK-----AHLKRAILGOBEALRIHALCRVIRE-----VDLLRAV 492
DB 331 EVERITTADYGASDEASTSIANKLAHLKEALASIEBAQELQRMCFELKNKASLTYMKEV 390
QY 493 ISQTLQRSIAKY-----AELDREDDFCEAAEAPDIQPKTHQKPEARMPLRSQKGPDIHF 547
DB 391 HSKS-QEELEKFRLOQASAKNIEDGDLERS-----VEASSKFTLSUTGPDFTSHKVDPFK 445
QY 548 RLGPUSVFSAKNRWRLVGPVHLTRG-----EGG----- 575
DB 446 RLGPFAIFSARRHWTAPRCVRLQKSSLYHVSFNDKCPDLNDDEHDGGVNYKKEEF 505
QY 576 --FGLTLRGDSPVLIAAIVIPGSOAAAAGLKEGDYIVSVNGQPCRWHRHAEVVTTELKAAGE 633
DB 506 ENFGFHVGDAPVITAHVEINSLADLGGIKEGDFIVEIAGVDVKWYSHQVQVQLIQSGS 565
QY 634 AGASLQVWSLLPSSRLPSLGRRPVLLGPRGLRSQREHGCKT--PASTWASPRPLNWS 691
DB 566 T-LELRVITPMDRNYLKPILSSK-----GSLSTLSAASSSGISSGFPSTIAAKPKLHL- 618
QY 692 RXAQOQKTGCGPQPCAPV 709
DB 619 -----KTSSSRPAGSV 630

Search completed: July 13, 2004, 12:04:25

Job time : 91 secs

mis Page Blank (uspto)

QY	121	CTGAGAGAGTGTGTGCTGGAGAGGGTGTCTGGTCCACAGAGGGGAAACCCAGTGGCT	180	QY	1201	GGCGAGTACTTCGGCTCCCTGGCCACCTAGCTAGCCATGGCCCTCTGGCGAGCTCC	1260
Db	223	CTGAGAGAGTGTGTGCTGGAGAGGGTGTCTGGTCCACAGAGGGGAAACCCAGTGGCT	282	Db	1303	GGCGAGTACTTCGGCTCCCTGGCCACCTAGCTAGCCATGGCCCTCTGGCGAGCTCC	1362
QY	181	GTGACGTTGGCCCATGATCGGACGCTTCATGCACTGGGCTGTCACTCCCTGAGCGCAG	240	QY	1261	CCAGCGACCGAGGAGAGCTCCCGACGACGAGAGGTCTTCTCTGAGAGCCGCCCACTCC	1320
Db	283	GTGACGTTGGCCCATGATCGGACGCTTCATGCACTGGGCTGTGACTCCCTGAGCGCAG	342	Db	1363	CCAGCGACCGAGGAGAGCTCCCGACGACGAGAGGTCTTCTCTGAGAGCCGCCCACTCC	1422
QY	241	ATCCAGTGGCGGCAGCTGCAGAGCCCGCAGGGCCAGATTACACAGCAGATTGACAGGAG	300	QY	1321	TCTAGCCCCGAGGCGCTGTCTGCGCAGAGGAGCTGAGAGCGCAGGAGCTTGGCAAG	1380
Db	343	ATCCAGTGGCGGCAGCTGCAGAGCCCGCAGGGCCAGATTACACAGCAGATTGACAGGAG	402	Db	1423	TCTAGCCCCGAGGCGCTGTCTGCGCAGAGGAGCTGAGAGCGCAGGAGCTTGGCAAG	1482
QY	301	CTGACATGCGGACGGGCGCTGAGAACCTTCTACAGAGCGACGAGAACAAACCGGGTGAGA	360	QY	1381	GCACACCTCAAGCGTGCCCATCTCTGGGCGCAGAGGAGCGCTGCGGTGACGCGCTGTGTC	1440
Db	403	CTGACATGCGGACGGGCGCTGAGAACCTTCTACAGAGCGACGAGAACAAACCGGGTGAGA	462	Db	1483	GCACACCTCAAGCGTGCCCATCTCTGGGCGCAGAGGAGCGCTGCGGTGACGCGCTGTGTC	1542
QY	361	GAGACGGTGGCCCTGGAGCTGAGCTAGCTCAACTCCAACTGCAGCTGTCTGAAGGAGGAG	420	QY	1441	CGGCTCTGCGGAGGTGACCTGTCTGGGCTGTGATCTCTCCAGAGGCTGACGCGCTCA	1500
Db	463	GAGACGGTGGCCCTGGAGCTGAGCTAGCTCAACTCCAACTGCAGCTGTCTGAAGGAGGAG	522	Db	1543	CGGCTCTGCGGAGGTGACCTGTCTGGGCTGTGATCTCTCCAGAGGCTGACGCGCTCA	1602
QY	421	CTGAGAGAGCTCAGCGGTGGCGTGGACCTCTGGCGGCGATGGAGCGAGCTGCACTGTC	480	QY	1501	CTGGCCCAAGTATGCGGAGCTCGAACCTGAGGATGACTTCTGTGAGGCTGCGGAGGCCG	1560
Db	523	CTGAGAGAGCTCAGCGGTGGCGTGGACCTCTGGCGGCGATGGAGCGAGCTGCACTGTC	582	Db	1603	CTGGCCCAAGTATGCGGAGCTCGAACCTGAGGATGACTTCTGTGAGGCTGCGGAGGCCG	1662
QY	481	CCCATGATCCCTTGGGCTGAGAGGAGACCAAGAGCTGACTGCTACACCGCTGAAG	540	QY	1561	GACATCCAGCCTAAGACCCACCAAGAGCCAGAGGCGCAGGATGCCACGCTGTCTCCAGGG	1620
Db	583	CCCATGATCCCTTGGGCTGAGAGGAGACCAAGAGCTGACTGCTACACCGCTGAAG	642	Db	1663	GACATCCAGCCTAAGACCCACCAAGAGCCAGAGGCGCAGGATGCCACGCTGTCTCCAGGG	1722
QY	541	GAGCTGATCTCAGTGACCTTTGGAGAGGAGCGGCGCTCTACAGAGGCGAGAAATCAGGGAG	600	QY	1621	AAGGGGCTGACATCTTCATCTGGCTGGGCGCCCTCTCTGTGTTCTCAGCAAGAACCGG	1680
Db	643	GAGCTGATCTCAGTGACCTTTGGAGAGGAGCGGCGCTCTACAGAGGCGAGAAATCAGGGAG	702	Db	1723	AAGGGGCTGACATCTTCATCTGGCTGGGCGCCCTCTCTGTGTCTCAGCAAGAACCGG	1782
QY	601	CTGAGAGCCCTGCGGACGGCCATCGGACCCCGACCGGAAATGAGTGGGCTGAGAGCTG	660	QY	1681	TGGCGGCTGTGGGCGCGCTCCACCTGACCCGAGGAGAGGGCGGCTTTGGCCCTCAAGCTT	1740
Db	703	CTGAGAGCCCTGCGGACGGCCATCGGACCCCGACCGGAAATGAGTGGGCTGAGAGCTG	762	Db	1783	TGGCGGCTGTGGGCGCGCTCCACCTGACCCGAGGAGAGGGCGGCTTTGGCCCTCAAGCTT	1842
QY	661	CTCAGAGCTATTACAAACAGCTGTGCTTCTGATGCGGCTTCTTCAACCCCTGCCAGG	720	QY	1741	CGGCGGAGACTCGGCTGTCTCTCATCTGCTGCGTCAATTCAGGAGGCGAGGCGGCGCT	1800
Db	763	CTCAGAGCTATTACAAACAGCTGTGCTTCTGATGCGGCTTCTTCAACCCCTGCCAGG	822	Db	1843	CGGCGGAGACTCGGCTGTCTCTCATCTGCTGCGTCAATTCAGGAGGCGAGGCGGCGCT	1902
QY	721	AGCCTCGGCTTCTTCCACTGTGTAGCTGCTTACTTGGGTCCTGGGTCCTGGGCGCCAGCGT	780	QY	1801	GGCTTGAAGAGGGCGACTACATTTGTGTGATGAGTGGGCGAGCCATCAGTGTGTGAGA	1860
Db	823	AGCCTCGGCTTCTTCCACTGTGTAGCTGCTTACTTGGGTCCTGGGTCCTGGGCGCCAGCGT	882	Db	1903	GGCTTGAAGAGGGCGACTACATTTGTGTGATGAGTGGGCGAGCCATCAGTGTGTGAGA	1962
QY	781	GCCCTGGCTTCGAGAGAGGGCAGGCTTCTTCTCAACATGCGTGCCTTCCACACGAGATT	840	QY	1861	CACGCGAGGTGGTGACGAGCTGAAGGCTGCGGAGAGGCGGCGCGCCAGCTTCAGGTG	1920
Db	883	GCCCTGGCTTCGAGAGAGGGCAGGCTTCTTCTCAACATGCGTGCCTTCCACACGAGATT	942	Db	1963	CACGCGAGGTGGTGACGAGCTGAAGGCTGCGGAGAGGCGGCGCGCCAGCTTCAGGTG	2022
QY	841	GGGCGCGCCAGGACCGCTCTGACCGAGGTCGCGCGCGCTATGGAGGCTTCCAG	900	QY	1921	GTGTGCTGTGCGCCAGCTCTAGACTGCGCCAGCTTGGGGGAGCGCGCGCGCTCTGCTG	1980
Db	943	GGGCGCGCCAGGACCGCTCTGACCGAGGTCGCGCGCGCTATGGAGGCTTCCAG	1002	Db	2023	GTGTGCTGTGCGCCAGCTCTAGACTGCGCCAGCTTGGGGGAGCGCGCGCGCTCTGCTG	2082
QY	901	AGGGCGCTGGGCTTCAGGCTCTCTGAGGAGAACTTCTCCCATGCGCGCGAGCCAGAC	960	QY	1981	GGCCCGCAGGGGCTTCTAAGGAGCCAGAGGAGCATGCTTGAAGACCCCGGCTCCACG	2040
Db	1003	AGGGCGCTGGGCTTCAGGCTCTCTGAGGAGAACTTCTCCCATGCGCGCGAGCCAGAC	1062	Db	2083	GGCCCGCAGGGGCTTCTAAGGAGCCAGAGGAGCATGCTTGAAGACCCCGGCTCCACG	2142
QY	961	ATGAGCGCTGCGTCCCTCTGCGCACTGGAGCAGCTCATGATGGCCCGAGGCCAGGAATGT	1020	QY	2041	TGGGCGCAGTCCCGGCGCCCTCTCTCAACTGAGGCGGAAAGGCCAGAGGGGCAAGACTGGA	2100
Db	1063	ATGAGCGCTGCGTCCCTCTGCGCACTGGAGCAGCTCATGATGGCCCGAGGCCAGGAATGT	1122	Db	2143	TGGGCGCAGTCCCGGCGCCCTCTCTCAACTGAGGCGGAAAGGCCAGAGGGGCAAGACTGGA	2202
QY	1021	GTGTTTGAAGGCTTCAACACCTGCTCCATGGCCCCCAGACTGCTGGCCAGCTG	1080	QY	2101	GGCTGCCCGCCAGCGCTGTGCGCCAGTGAAGCAGCTCGGCGCTCATCTCTGAGCACCCA	2160
Db	1123	GTGTTTGAAGGCTTCAACACCTGCTCCATGGCCCCCAGACTGCTGGCCAGCTG	1182	Db	2203	GGCTGCCCGCCAGCGCTGTGCGCCAGTGAAGCAGCTCGGCGCTCATCTCTGAGCACCCA	2262
QY	1081	CGCCTGGCGCAGGAGCGCGCCAGGTGGCAGCCAGTACAGGCTAGTGCACCGGACCATG	1140	QY	2161	GGGTGGCGG 2169	
Db	1183	CGCCTGGCGCAGGAGCGCGCCAGGTGGCAGCCAGTACAGGCTAGTGCACCGGACCATG	1242	Db	2263	GGGTGGCGG 2271	
QY	1141	GCCAGCACCGCTCCAGCATTAAGTGGCTGTCTCTGAGCTGCGCTTGTGATGTCAAG	1200				
Db	1243	GCCAGCACCGCTCCAGCATTAAGTGGCTGTCTCTGAGCTGCGCTTGTGATGTCAAG	1302				

Db	1903	GGCCTGAAGGAGGGCGGACTACATTGTGTCTAGTCAATGGCGAGCCATGACAGTGGTGGAGA	1962
Qy	1861	CACGCGAGTGGTGACGGAGCTGAAAGCTGCCGGAGAGCGGCGCCAGCCTGCAGGTG	1920
Db	1963	CACGCGAGTGGTGACGGAGCTGAAAGCTGCCGGAGAGCGGCGCCAGCCTGCAGGTG	2022
Qy	1921	GTGTGCGTGTGCCACGCTCTAGACTGCCCAGGCTTGGGGGACCGCGGCCGCTCTGCTG	1980
Db	2023	GTGTGCGTGTGCCACGCTCTAGACTGCCCAGGCTTGGGGGACCGCGGCCGCTCTGCTG	2082
Qy	1981	GGCCCCAGGGGGCTTTAAAGAGCCAGAGGAGCATGTGGTGCAAAGACCCCGGATCCACG	2040
Db	2083	GGCCCCAGGGGGCTTTAAAGAGCCAGAGGAGCATGTGGTGCAAAGACCCCGGATCCACG	2142
Qy	2041	TGGGCGCAGTCCCGGGCCCTCTCAACTGGAGCCGAAAGCCACGACGAGGCAAGACTGGA	2100
Db	2143	TGGGCGCAGTCCCGGGCCCTCTCAACTGGAGCCGAAAGCCACGACGAGGCAAGACTGGA	2202
Qy	2101	GGCTGCCGCCCCAGCCGCTGTGCCCCAGTGAAGCCAGCTCCGCCCTCATCTTTGAAGCACCCA	2160
Db	2203	GGCTGCCGCCCCAGCCGCTGTGCCCCAGTGAAGCCAGCTCCGCCCTCATCTTTGAAGCACCCA	2262
Qy	2161	GGGTGGCGG	2169
Db	2263	GGGTGGCGG	2271

RESULT 3

```

US-10-451-207-11
; Sequence 11, Application US/10451207
; Publication No. US20040038267A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: BURFORD, Neil
; APPLICANT: DING, Li
; APPLICANT: YUE, Henry
; APPLICANT: THORNTON, Michael B.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: SWARNAKAR, Anita
; APPLICANT: DUGGAN, Brendan M.
; APPLICANT: LU, Dyrong Aina M.
; APPLICANT: THANGAVELU, Kavitha
; APPLICANT: WARREN, Bridget A.
; APPLICANT: TANG, Y. Tom
; APPLICANT: KHAN, Farrah A.
; APPLICANT: YAO, Monique G.
; APPLICANT: EMERLING, Brooke M.
; TITLE OF INVENTION: INTRACELLULAR SIGNALING MOLECULES
; FILE REFERENCE: PF-0868 USN
; CURRENT APPLICATION NUMBER: US/10/451,207
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: PCT/US01/50315
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 60/257,804
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 60/260,102
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 11
; LENGTH: 2837
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040038267A1 5560161CB1
US-10-451-207-11

Query Match      89.9%; Score 1951; DB 13; Length 2837;
Best Local Similarity 100.0%; Pred. No. 0;

```

```
QY 1299 CTTCTGCGAGCCGCCCACTCTCTAAAGCCCGAGGCCCTGTGCTCCCGAGAGAGCTGGA 1358
Db 1261 CTTCTGCGAGCCGCCCACTCTCTAAAGCCCGAGGCCCTGTGCTCCCGAGAGAGCTGGA 1320
QY 1359 GGAGCGCAGCAGCTTGGCAAGGCACACCTGAAGCGTGCATCTCGGGCAGGAGGAGGC 1418
Db 1321 GGAGCGCAGCAGCTTGGCAGGCACACCTGAAGCGTGCATCTCGGGCAGGAGGAGGC 1380
QY 1419 GCTGGCGCTGCGAGCCCTGTGCGCGCTCTGCGCGAGGTGACCTCTTCGGGCTGTGAT 1478
Db 1381 GCTGGCGCTGCGAGCCCTGTGCGCGCTCTGCGCGAGGTGACCTCTTCGGGCTGTGAT 1440
QY 1479 CTTCCAGACGCTGCGAGCGCTCACTGCCCAAGTATGCGGAGCTCGACCGTGAAGATGACTT 1538
Db 1441 CTTCCAGACGCTGCGAGCGCTCACTGCCCAAGTATGCGGAGCTCGACCGTGAAGATGACTT 1500
QY 1539 CTGTGAGGCTGCGAGGCCCGGACATCCAGGCTTAAGACCCAGAGGCCAGAGGCCAG 1598
Db 1501 CTGTGAGGCTGCGAGGCCCGGACATCCAGGCTTAAGACCCAGAGGCCAGAGGCCAG 1560
QY 1599 GATGCGACGCTGTCCAGGGGAGAGGGGCTGACATCTTCCATCGGCTGGGGCCCTGTGC 1658
Db 1561 GATGCGACGCTGTCCAGGGGAGAGGGGCTGACATCTTCCATCGGCTGGGGCCCTGTGC 1620
QY 1659 TGTGTTCTCAGCAAGAACCGGTGGCGGTGGTGGGGCCGCTCCACCTGACCCGAGAGA 1718
Db 1621 TGTGTTCTCAGCAAGAACCGGTGGCGGTGGTGGGGCCGCTCCACCTGACCCGAGAGA 1680
QY 1719 GGGCGGCTTTGGCTCAGCGCTTTCGGGGAGACTCGCTGTGCTCATCGCTGCGCTCATTC 1778
Db 1681 GGGCGGCTTTGGCTCAGCGCTTTCGGGGAGACTCGCTGTGCTCATCGCTGCGCTCATTC 1740
QY 1779 AGGAGCCAGCGCGCGGCGCTGGCTGGAAGAGGGGACTACATGTGTCAAGTGAATGG 1838
Db 1741 AGGAGCCAGCGCGCGGCGCTGGCTGGAAGAGGGGACTACATGTGTCAAGTGAATGG 1800
QY 1839 GCAGCCATGCAAGTGTGGAGACACGCGAGGTGGTGGAGGCTGAAGGCTGCGGGAGA 1898
Db 1801 GCAGCCATGCAAGTGTGGAGACACGCGAGGTGGTGGAGGCTGAAGGCTGCGGGAGA 1860
QY 1899 GCGCGGCGCCAGCTGCAAGTGTGTGCTGCGCTGCGCCAGCTCTAGACTGCCAGCTTGGG 1958
Db 1861 GCGCGGCGCCAGCTGCAAGTGTGTGCTGCGCTGCGCCAGCTCTAGACTGCCAGCTTGGG 1920
QY 1959 GGAACCGCGCGCGCTGCTGCTGGGGCCCAAGGGGCTTCTAAGGAGCCAGAGGAGCATGG 2018
Db 1921 GGAACCGCGCGCGCTGCTGCTGGGGCCCAAGGGGCTTCTAAGGAGCCAGAGGAGCATGG 1980
QY 2019 TTGAAGACCCCGGATCCAGCTGCGGCCAGTCCCGCGCCCTCTCTCACTGGAGCCGAAA 2078
Db 1981 TTGAAGACCCCGGATCCAGTGGGCGAGTCCCGCGCCCTCTCTCACTGGAGCCGAAA 2040
QY 2079 GGCACAGAGGCGAAGACTGGAGGCTGCCCCAGCCCTGTGCCCCAGTGAAGCCAGCTCC 2138
Db 2041 GGCACAGAGGCGAAGACTGGAGGCTGCCCCAGCCCTGTGCCCCAGTGAAGCCAGCTCC 2100
QY 2139 GCCCTCATCTTGAAGCACCCAGGGTGGCCG 2169
Db 2101 GCCCTCATCTTGAAGCACCCAGGGTGGCCG 2131
```

RESULT 4

```
US-10-112-944-74
; Sequence 74, Application US/10112944
; Publication No. US20040048249A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Gezhi
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
```

```
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Zhiwei
; TITLE OF INVENTION: No. US20040048249A1 Nucleic Acids and
; TITLE OF INVENTION: Secreted Polypeptides
; FILE REFERENCE: 805A
; CURRENT APPLICATION NUMBER: US/10/112,944
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 2000-01-21
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US 09/519,705
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 09/552,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 924
; SOFTWARE: pt_FL_genes Version 5.0
; SEQ ID NO 74
; LENGTH: 2738
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (103)..(2109)
; US-10-112-944-74
```

Query Match 89.2%; Score 1935; DB 13; Length 2738;
Best Local Similarity 99.8%; Pred. No. 0;

Matches 1948; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

```
QY 219 GGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGGCGAGCTGCAGAGCCGCGAGGCCAGAT 278
Db 162 GGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGGCGAGCTGCAGAGCCGCGAGGCCAGAT 221
QY 279 TCACGAGCAGATTGACAGAGGAGCTGCAGATCGGAGCGGCGCTGAGAACCTCTACAGAGC 338
Db 222 TCACGAGCAGATTGACAGAGGAGCTGCAGATCGGAGCGGCGCTGAGAACCTCTACAGAGC 281
QY 339 CACGAGCAACACCGGCTGAGAGAGACGCTCGCCCTGGAGCTGAGCTACGTCAACTCCAA 398
Db 282 CACGAGCAACACCGGCTGAGAGAGACGCTCGCCCTGGAGCTGAGCTACGTCAACTCCA 341
QY 399 CTTGAGCTGCTGAAGGAGGAGCTGGAGAGCTCAGCGGTGGCGCTGAGACCTTGGCCGCA 458
Db 342 CTTGAGCTGCTGAAGGAGGAGCTGGAGAGCTCAGCGGTGGCGCTGAGACCTTGGCCGCA 401
QY 459 TGGGAGCGAAGCTGTCACTGTCCCATGATCCCCCTGGGCTTGAAGAGACCAAGGAGCT 518
Db 402 TGGGAGCGAAGCTGTCACTGTCCCATGATCCCCCTGGGCTTGAAGAGACCAAGGAGCT 461
QY 519 GGACTGCTTACACCGCTGAAGGAGCTGATCTCAGTGCACTTTTGGAGAGGACGCGGCTC 578
Db 462 GGACTGCTTACACCGCTGAAGGAGCTGATCTCAGTGCACTTTTGGAGAGGACGCGGCTC 521
QY 579 CTACGAGCGAAGATCAAGGAGCTGGAGGCCCTTGGCGAGGCCCATCGGACCCCGCCG 638
Db 522 CTACGAGCGAAGATCAAGGAGCTGGAGGCCCTTGGCGAGGCCCATCGGACCCCGCCG 581
QY 639 GAATGAGTCCGGCTGGAGCTGCTCAGAGCTATTACACAGCTGTGCTTCTTGATGTC 698
Db 582 GAATGAGTCCGGCTGGAGCTGCTCAGAGCTATTACACAGCTGTGCTTCTTGATGTC 641
```


QY 1615 CAGGGAGGGGCTGACATCTTCCATCGGCTGGGGCCCTGCTCTGTGTCTCAGCCAG 1674
Db |||||||
QY 373 CAGGGAGGGGCTGACATCTTCCATCGGCTCCGGCCCTGCTGTGTCTCAGCCAG 432
Db |||||||
QY 1675 AACCGGTGGCGGTGTGTGGGGCCGTCCACCTGACCCGAGGAGGCGGCTTTGGCCCTC 1734
Db |||||||
QY 433 AACCGGTGGCGGTGTGTGGGGCCGTCCACCTGACCCGAGGAGGCGGCTTTGGCCCTC 492
QY 1735 AGCTTTGGGGAGACTCGGCTGTCTCATCGCTGCGCTCATTCAGGAGCCAGGCGCG 1794
Db |||||||
QY 493 AGCTTTGGGGAGACTCGGCTGTCTCATCGCTGCGCTCATTCAGGAGCCAGGCGCG 552
QY 1795 GCGGCTGGCTTGAAGGAGGCGGCTGACATCTTGTGTGAGTGAATGGGAGGCGGCTG 1854
Db |||||||
QY 553 GCGGCTGGCTTGAAGGAGGCGGCTGACATCTTGTGTGAGTGAATGGGAGGCGGCTG 612
QY 1855 TGGAGACACGCGGAGGTGTGTGAGGAGCTGAAGGCTGCGGAGGAGGCGGCGGCGGCTG 1914
Db |||||||
QY 613 TGGAGACACGCGGAGGTGTGTGAGGAGCTGAAGGCTGCGGAGGAGGCGGCGGCGGCTG 672
QY 1915 CAGGTGGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1974
Db |||||||
QY 673 CAGGTGGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 732
QY 1975 CTGCTGGGCGCCAGGGGGCTTCTAAGGAGCCAGAGGAGCATGTTGCAAGACCCCGCA 2034
Db |||||||
QY 733 CTGCTGGGCGCCAGGGGGCTTCTAAGGAGCCAGAGGAGCATGTTGCAAGACCCCGCA 792
QY 2035 TCCAGTGGGCGAGTCCCGGCGCCCTCTCAACTGGAGCCGAAAGGCCCCAGAGGGCAAG 2094
Db |||||||
QY 793 TCCAGTGGGCGAGTCCCGGCGCCCTCTCAACTGGAGCCGAAAGGCCCCAGAGGGCAAG 852
QY 2095 ACTGAGGCTGCCCCAGGCGCTGTGCCCCAGTGAAGCCAGTCCGCGCTCATCTTTGAAG 2154
Db |||||||
QY 853 ACTGAGGCTGCCCCAGGCGCTGTGCCCCAGTGAAGCCAGTCCGCGCTCATCTTTGAAG 912
QY 2155 CACCCAGGCTGGCCG 2169
Db |||||||

RESULT 6

US-10-023-896-15
; Sequence 15, Application US/10023896
; Publication No. US2003002776A1
; GENERAL INFORMATION:
; APPLICANT: Victor Roschke
; TITLE OF INVENTION: 29 Human Cancer Associated Proteins
; FILE REFERENCE: PA004PI
; CURRENT APPLICATION NUMBER: US/10/023,896
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: PCT/US00/23794
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/152,296
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/158,003
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-896-15

Query Match 36.0%; Score 781.4; DB 15; Length 1013;
Best Local Similarity 99.7%; Pred. No. 3.6e-181;
Matches 793; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1375 GSCAAGGCACACTCAAGGTGCCATCTCTGGGGAGGAGGCGCTGGGCTGCACGCC 1434
|||

Db 27 GGCAGGCGACACTGAAGCGTGCATCTCTGGGGCAGGAGAGCGCTGGGCTGCACGCC 86
QY 1435 CTGTGCGCGCTCTCTGGCGAGGTGGACCTGTCTTGGGGCTGTGATCTCCAGAGCGCTGCAG 1494
Db |||||||
QY 87 CTGTGCGCGCTCTCTGGCGAGGTGGACCTGTCTTGGGGCTGTGATCTCCAGAGCGCTGCAG 146
QY 1495 CGTCTACTGGCGCAAGTATGCGAGGTGCGAGCTGACCTTCTGTGTGAGGTGCGGAG 1554
Db |||||||
QY 147 CGTCTACTGGCGCAAGTATGCGAGGTGCGAGCTGACCTTCTGTGTGAGGTGCGGAG 206
QY 1555 GCGCGGACATCCAGCCTAAGACCCACAGAGCCAGAGCCAGAGTGCACGCGCTGTCC 1614
Db |||||||
QY 207 GCGCGGACATCCAGCCTAAGACCCACAGAGCCAGAGCCAGAGTGCACGCGCTGTCC 266
QY 1615 CAGGGGAAAGGGGCTTGCATCTTCCATCGGCTGGGGCCCTCTGTGTGTCTCAGCAAG 1674
Db |||||||
QY 267 CAGGGGAAAGGGGCTTGCATCTTCCATCGGCTGGGGCCCTCTGTGTGTCTCAGCAAG 326
QY 1675 AACCGGTGGCGCTGTGTGGGGCCCTGCTCACTGACCCGAGGAGAGGGCGCTTTGGCCCTC 1734
Db |||||||
QY 327 AACCGGTGGCGCTGTGTGGGGCCCTGCTCACTGACCCGAGGAGAGGGCGCTTTGGCCCTC 386
QY 1735 AGCTTTGGGGAGACTCGGCTGTCTCATCGCTGCGGTCAATTCAGGAGCCAGGCGCG 1794
Db |||||||
QY 387 AGCTTTGGGGAGACTCGGCTGTCTCATCGCTGCGGTCAATTCAGGAGCCAGGCGCG 446
QY 1795 GCGGCTGGGCTTGAAGGAGGGCGGCTACATTTGTGTGAGTGAATGGGCGAGGCTG 1854
Db |||||||
QY 447 GCGGCTGGGCTTGAAGGAGGGCGGCTACATTTGTGTGAGTGAATGGGCGAGGCTG 506
QY 1855 TGGAGACACGCGGAGGTGTGTGCGAGGTGAAGGCTGCGGGAGAGCGGCGCGCAGCCCTG 1914
Db |||||||
QY 507 TGGAGACACGCGGAGGTGTGTGCGAGGTGAAGGCTGCGGGAGAGCGGCGCGCAGCCCTG 566
QY 1915 CAGGTGGTGTGCTGCTGCCAGCTTAGACTGCGGAGCTGGGGAGAGCGGCGCGCCCTC 1974
Db |||||||
QY 567 CAGGTGGTGTGCTGCTGCCAGCTTAGACTGCGGAGCTGGGGAGAGCGGCGCGCCCTC 626
QY 1975 CTGCTGGGCGCCAGGGGGCTTCTAAGGAGCCAGAGGAGCATGTTGCAAGACCCCGCA 2034
Db |||||||
QY 627 CTGCTGGGCGCCAGGGGGCTTCTAAGGAGCCAGAGGAGCATGTTGCAAGACCCCGCA 686
QY 2035 TCCAGTGGGCGAGTCCCGGCGCCCTCTCTCAACTGAGCCGAAAGGCCAGAGGCGCAAG 2094
Db |||||||
QY 687 TCCAGTGGGCGAGTCCCGGCGCCCTCTCTCAACTGAGCCGAAAGGCCAGAGGCGCAAG 746
QY 2095 ACTGAGGCTGCCCGGCGGCTGTGCCCCAGTGAAGCCAGTCCGCGCTCATCTTTGAAG 2154
Db |||||||
QY 747 ACTGAGGCTG-CCCCAGCTGTGCCAGTGAAGCCAGTCCGCGCTCATCTTTGAAG 805
QY 2155 CACCCAGGCTGGCCG 2169
Db |||||||
QY 806 CACCCAGGCTGGCCG 820
Db |||||||

RESULT 7

US-09-925-301-174
; Sequence 174, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 174
; LENGTH: 1013

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-301-174

Query Match      36.0%; Score 779.8; DB 9; Length 1013;
Best Local Similarity 99.6%; Pred. No. 8.8e-181;
Matches 792; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1375 GGCAGGACACACCTGAAGCGTGCATCTCTGGGGCAGGAGGCGCTGCGGCTGCAGGCC 1434
Db      27 GGCAGGACACACCTGAAGCGTGCATCTCTGGGGCAGGAGGCGCTGCGGCTGCAGGCC 86

QY 1435 CTGTGCGCGCTCTCGCGAGGTGGACCTCTTGGGCTGTGATCTCCAGACGCTCGAG 1494
Db      87 CTGTGCGCGCTCTCGCGAGGTGGACCTCTTGGGCTGTGATCTCCAGACGCTCGAG 146

QY 1495 CGCTCACTGCGCAAGTATGCGGAGCTCGACCGTGAAGATGACTTCTGTGAGGCTGCCGAG 1554
Db      147 CGCTCACTGCGCAAGTATGCGGAGCTCGACCGTGAAGATGACTTCTGTGAGGCTGCCGAG 206

QY 1555 GCCCGGACATCCAGCTTAAGACCCAGAGCCAGAGCCAGGATGCCAGCTGTCC 1614
Db      207 GCCCGGACATCCAGCTTAAGACCCAGAGCCAGAGCCAGGATGCCAGCTGTCC 266

QY 1615 CAGGGAGGGGCTGCATCTTCCATCGGCTGGGCGCCCTGTCTGTGTTCTCAGCCAAAG 1674
Db      267 CAGGGAGGGGCTGCATCTTCCATCGGCTGGGCGCCCTGTCTGTGTTCTCAGCCAAAG 326

QY 1675 AACCGGTGGGGGCTGGTGAAGCGAGCTGAAGCGCTCGGAGAGGCGGCGCCAGCTTG 1734
Db      327 AACCGGTGGGGGCTGGTGAAGCGAGCTGAAGCGCTCGGAGAGGCGGCGCCAGCTTG 386

QY 1735 ACGTGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACCGCGCCGCTC 1794
Db      387 ACGTGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACCGCGCCGCTC 446

QY 1795 GCGGCTGCGCTGAAGGAGGCGACTACATTGTGTGATGTAATGGGAGCCATGCAAGTGG 1854
Db      447 GCGGCTGCGCTGAAGGAGGCGACTACATTGTGTGATGTAATGGGAGCCATGCAAGTGG 506

QY 1855 TGGAGACACGCGGAGGTGGTGAAGCGAGCTGAAGCGCTCGGAGAGGCGGCGCCAGCTTG 1914
Db      507 TGGAGACACGCGGAGGTGGTGAAGCGAGCTGAAGCGCTCGGAGAGGCGGCGCCAGCTTG 566

QY 1915 CAGTGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACCGCGCCGCTC 1974
Db      567 CAGTGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACCGCGCCGCTC 626

QY 1975 CTGCTGGGCGCCAGGGGCTTCTAAGAGCCAGAGGAGCATGTTGTGCAAGACCCCGGCA 2034
Db      627 CTGCTGGGCGCCAGGGGCTTCTAAGAGCCAGAGGAGCATGTTGTGCAAGACCCCGGCA 686

QY 2035 TCCACGTGGGCGAGTCCCGGCGCCCTCTCTCAACTGGAGCCGAAAGGCCAGCGGCAAG 2094
Db      687 TCCACGTGGGCGAGTCCCGGCGCCCTCTCTCAACTGGAGCCGAAAGGCCAGCGGCAAG 746

QY 2095 ACTGGAGGCTGCCCGCCAGCCCTGTGCCCCAGTGAAGCCAGCTCCGCTCTATCTTTGAG 2154
Db      747 ACTGGAGGCTG-CCCCAGCCCTGTGCCCCAGTGAAGCCAGCTCCGCTCTATCTTTGAG 805

2155 CACCCAGGGTGGCCG 2169
806 CACCCAGGGTGGCCG 820
```

```
RESULT 8
US-10-023-896-43
; Sequence 43, Application US/10023896
; Publication No. US2003002776A1
; GENERAL INFORMATION:
; APPLICANT: Victor Roschke
; TITLE OF INVENTION: 29 Human Cancer Associated Proteins
; FILE REFERENCE: PA004P1
```

```
; CURRENT APPLICATION NUMBER: US/10/023,896
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: PCT/US00/23794
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/152,296
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/158,003
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 1013
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-896-43

Query Match      36.0%; Score 779.8; DB 15; Length 1013;
Best Local Similarity 99.6%; Pred. No. 8.8e-181;
Matches 792; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1375 GGCAGGACACACCTGAAGCGTGCATCTCTGGGGCAGGAGGCGCTGCGGCTGCAGGCC 1434
Db      27 GGCAGGACACACCTGAAGCGTGCATCTCTGGGGCAGGAGGCGCTGCGGCTGCAGGCC 86

QY 1435 CTGTGCGCGCTCTCGCGAGGTGGACCTCTTGGGCTGTGATCTCCAGACGCTCGAG 1494
Db      87 CTGTGCGCGCTCTCGCGAGGTGGACCTCTTGGGCTGTGATCTCCAGACGCTCGAG 146

QY 1495 CGCTCACTGCGCAAGTATGCGGAGCTCGACCGTGAAGATGACTTCTGTGAGGCTGCCGAG 1554
Db      147 CGCTCACTGCGCAAGTATGCGGAGCTCGACCGTGAAGATGACTTCTGTGAGGCTGCCGAG 206

QY 1555 GCCCGGACATCCAGCTTAAGACCCAGAGCCAGAGCCAGGATGCCAGCTGTCC 1614
Db      207 GCCCGGACATCCAGCTTAAGACCCAGAGCCAGAGCCAGGATGCCAGCTGTCC 266

QY 1615 CAGGGAGGGGCTGCATCTTCCATCGGCTGGGCGCCCTGTCTGTGTTCTCAGCCAAAG 1674
Db      267 CAGGGAGGGGCTGCATCTTCCATCGGCTGGGCGCCCTGTCTGTGTTCTCAGCCAAAG 326

QY 1675 AACCGGTGGGGGCTGGTGAAGCGAGCTGAAGCGCTCGGAGAGGCGGCGCCAGCTTG 1734
Db      327 AACCGGTGGGGGCTGGTGAAGCGAGCTGAAGCGCTCGGAGAGGCGGCGCCAGCTTG 386

QY 1735 ACGTGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACCGCGCCGCTC 1794
Db      387 ACGTGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACCGCGCCGCTC 446

QY 1795 GCGGCTGCGCTGAAGGAGGCGACTACATTGTGTGATGTAATGGGAGCCATGCAAGTGG 1854
Db      447 GCGGCTGCGCTGAAGGAGGCGACTACATTGTGTGATGTAATGGGAGCCATGCAAGTGG 506

QY 1855 TGGAGACACGCGGAGGTGGTGAAGCGAGCTGAAGCGCTCGGAGAGGCGGCGCCAGCTTG 1914
Db      507 TGGAGACACGCGGAGGTGGTGAAGCGAGCTGAAGCGCTCGGAGAGGCGGCGCCAGCTTG 566

QY 1915 CAGTGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACCGCGCCGCTC 1974
Db      567 CAGTGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACCGCGCCGCTC 626

QY 1975 CTGCTGGGCGCCAGGGGCTTCTAAGAGCCAGAGGAGCATGTTGTGCAAGACCCCGGCA 2034
Db      627 CTGCTGGGCGCCAGGGGCTTCTAAGAGCCAGAGGAGCATGTTGTGCAAGACCCCGGCA 686

QY 2035 TCCACGTGGGCGAGTCCCGGCGCCCTCTCTCAACTGGAGCCGAAAGGCCAGCGGCAAG 2094
Db      687 TCCACGTGGGCGAGTCCCGGCGCCCTCTCTCAACTGGAGCCGAAAGGCCAGCGGCAAG 746

QY 2095 ACTGGAGGCTGCCCGCCAGCCCTGTGCCCCAGTGAAGCCAGCTCCGCTCTATCTTTGAG 2154
Db      747 ACTGGAGGCTG-CCCCAGCCCTGTGCCCCAGTGAAGCCAGCTCCGCTCTATCTTTGAG 805
```



```
QY 2155 CACCCAGGTTGGCCG 2169
Db 806 CACCCAGGTTGGCCG 820

RESULT 9
US-10-112-944-558/c
; Sequence 558, Application US/10112944
; Publication No. US20040048249A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Gezhi
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Zhiwei
; TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
; TITLE OF INVENTION: Secreted Polypeptides
; FILE REFERENCE: 805A
; CURRENT APPLICATION NUMBER: US/10/112,944
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US 09/519,705
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 09/552,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 924
; SOFTWARE: pt_FL_genes Version 5.0
; SEQ ID NO 558
; LENGTH: 1046
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(1046)
; OTHER INFORMATION: n = a,t,c or g
US-10-112-944-558

Query Match 24.2%; Score 524; DB 13; Length 1046;
Best Local Similarity 99.1%; Pred. No. 2.7e-118;
Matches 527; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1638 CCATCGGCTGGGGCCCTCTGTGTCTTCAGCCAGAACCCGTTGGCGGCTGGTGGGGCC 1697
Db 702 CCCCTGCCAGGGGCCCTCTGTGTCTTCAGCCAGAACCCGTTGGCGGCTGGTGGGGCC 643

QY 1698 CGTCCACCTGACCCCGAGGAGAGGGCGCTTTGGCCCTCAGCTTCGSGGAGACTCGCCTGT 1757
Db 642 CGTCCACCTGACCCCGAGGAGAGGGCGGCTTTGGCCCTCAGCTTCGSGGAGACTCGCCTGT 583

QY 1758 CTTCTATCGCTGCGTCAATTCCAGGGAGCCAGGCGCGGGCTGGCTTGAAGAGGGCGA 1817
Db 582 CTTCTATCGCTGCGTCAATTCCAGGGAGCCAGGCGCGGGCTGGCTTGAAGAGGGCGA 523

QY 1818 CTACATTGTTCAGTGAATGGGAGCCATGTCAGGTGGTGAGACACCGCGAGGTGGTGAC 1877
Db 522 CTACATTGTTCAGTGAATGGGAGCCATGTCAGGTGGTGAGACACCGCGAGGTGGTGAC 463
QY 1878 GGAGCTGAAGGCTGCGGGAGAGCGGGCGCCAGCCTGCAAGGTGGTGTGCTGCTGCCCCAG 1937
Db 462 GGAGCTGAAGGCTGCGGGAGAGCGGGCGCCAGCCTGCAAGGTGGTGTGCTGCTGCCCCAG 403
QY 1938 CTCTAGACTGCCAGCTTTGGGGAGCCGCGGGCCCTCTGTGTGGGCCCCAGGGGGCTTCT 1997
Db 402 CTCTAGACTGCCAGCTTTGGGGAGCCGCGGGCCCTCTGTGTGGGCCCCAGGGGGCTTCT 343
QY 1998 AAGGAGCCAGAGGAGCATGTTGCAAGACCCCGGCAATCCAGTGGGCCAGTCCCGGGCC 2057
Db 342 AAGGAGCCAGAGGAGCATGTTGCAAGACCCCGGCAATCCAGTGGGCCAGTCCCGGGCC 283
QY 2058 CCTCTCAACTGGAGCCGAAAGGCCAGCAGGAGCAAGACTGGAGGCTGCCCGCCCTG 2117
Db 282 CCTCTCAACTGGAGCCGAAAGGCCAGCAGGAGCAAGACTGGAGGCTGCCCGCCCTG 223
QY 2118 TCCCCAGTGAAGCCAGCTCCGCCCTCATCTTGAAGCACCCAGGCTGGCCG 2169
Db 222 TGCCCCAGTGAAGCCAGCTCCGCCCTCATCTTGAAGCACCCAGGCTGGCCG 171

RESULT 10
US-10-239-607-52
; Sequence 52, Application US/10239607
; Publication No. US20030219761A1
; GENERAL INFORMATION:
; APPLICANT: Saatioglu, Fahri
; TITLE OF INVENTION: No. US20030219761A1el Prostate-Specific or
; TITLE OF INVENTION: Testis-Specific Nucleic Acid Molecules, Polypeptides, and
; TITLE OF INVENTION: Diagnostic and Therapeutic Methods
; FILE REFERENCE: 50218/003002
; CURRENT APPLICATION NUMBER: US/10/239,607
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09410
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,929
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 2100
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-239-607-52

Query Match 22.9%; Score 496.4; DB 16; Length 2100;
Best Local Similarity 56.8%; Pred. No. 1.5e-111;
Matches 977; Conservative 0; Mismatches 731; Indels 12; Gaps 3;

QY 219 GGGCTGTGACTCCCTGACGACGATCCAGTGGCGGCAGCTGCAGAGCCGAGGCGCCAGAT 278
Db 69 GGGCTGTGATCCCTTGCACAAACCGCGGAGTAAATTCAGAAATCAAGAGCTGCTTT 128
QY 279 TCACCAGCAGATTCACAGAGAGCTGCAGATGGGCGGCGCTGCAGAACTCTACAGAGC 338
Db 129 GAATCAGCAGATTCCTGAAAGCCGTCGGATGAGGACCGGAGCGGAAACCTTCTGAAAGT 188
QY 339 CACGAGCAACACCGGCTGAGAGAGACCGTGCCTCTGAGCTGAGCTAGCTCAACTCCA 398
Db 189 GGCACAAACTCAAAGGTGCGGGAGCAAGTGGGCTGGAGCTGAGCTTCTGCAACTCAGA 248
QY 399 CCTGAGCTGCTGAAGGAGGAGCTGCAGAGAGCTCAGCGTGGCGTGGACCTTGGCGGCA 458
Db 249 CTTGAGATGCTCAAGGAGAGCTGCAGAGGCTGAACATCTCGTGGGCTTATCAGAA 308
QY 459 TGGAGCGAAGAGTGTCACTGTCCCATGATCCCCCTGGGCTTGAAGAGAGCAAGAGAGCT 518
Db 309 CACAGAGGAGGATTTACGATTTCCCTGATTCCTTGGCTGAGAGGAAACGAAAGAGCT 368
QY 519 GGACTGGTCTACACCGCTGAAGGAGCTGATCTCAGTGCACTTTTGGAGAGGAGCGGCGCTC 578
```


||||| 369 CGACTTTGCGAGTCTCTCAGGATTTATCTCGAACAATTACAGTGAAGATGCGTATTT 428
||||| 579 CTACGAGCAGAAATCAGGAGCTGAGGCCCTCGCGCAGGCCATGCGGACCCCGACCG 638
||||| 429 ATATGAAGATGAAATTTGAGATCTTATGATCTGAGACAAGCTTTGTCGAGCGCTAGCG 488
||||| 639 GAATGAGTCCGGCTCGAGCTGCTCAGACCTATTTACAAACAGCTGCTCTTCTTGATGC 598
||||| 489 GGATGAGCCCGGGTGGAACTGCTGATGACATATCTCATCCAGCTGGCTTTTTCGAGAG 548
||||| 699 GCGCTCTCTCAACCTCGCCAGGAGCCTCGGCTCTTTCTTCACCTGGTACGACTCGCTTAC 758
||||| 549 TCGATTTCTCCGCGCCACACGGCAGATGGGACTCTCTTTCACCTGGTATGACTCTCTCAC 508
||||| 759 TGGGGTCCCGCCAGCAGCTGCTGCGCTTCGAGAAAGGCGAGCTTCTCTTCAACAT 818
||||| 609 TGGGGTTCCGGTTCAGCCAGCAGAACTCTGCTGCGAGAAAGCCAGTGTCTGTTCACAC 668
||||| 819 CGGTGCCCTCCACACGAGATTTGGGGCGCGCAGACCGCTCTGACCGAGGGTGC 878
||||| 669 TGGGGCTCTACACCCAGATTTGGGACCCGGTGGATCGGCAGACGGCTGGCTGGGA 728
||||| 879 CCGCGCTATGAGGCTTTCCAGAGGGCGCTGGGGCTTTACGCTCTCTGAGGAGAACTT 938
||||| 729 GAGTGCATAGATGCTTTTCAGAGAGCGCAGAGGGTTTAAATTTACCTGAAAGACACATT 788
||||| 939 CTCCTATGCGCGCAGCCAGACATGAGCGCTGCTGCTCTGCGCAGCTGGAGAGCTCAT 998
||||| 789 TACCACATCTCAAGTTACGACATGAGCCCTGCCATGCTCAGCGTGTCTGTCGTAAGATGAT 848
||||| 999 GATGGCCAGGCCACAGGAATGTGTTTGGGGCTCTCACACCTGCTCCATGGCCCC 1058
||||| 849 GCTTGCACAGCCCAAGAAAGCGTGTGGAAATCAGCCTTCTG-----GATCCG 902
||||| 1059 CCAAGACTGCTGCGCCAGCTGCGCTGGCGCAGAGCGCCCGCAGGTGGCAGCGAGTA 1118
||||| 903 GAATGAATTTCTATGCTGTGTGAAGTGGCTCAGAGGCTCTAAGGTGGGAGAGTCTA 962
||||| 1119 CAGCTAGTGACCGGACCATGGCCAGCAGCCCGTCCACGACTACGTGCTGTCTCTG 1178
||||| 963 CCAACAGCTACACGAGCAGTGGAGCCAGGCGCGGTGAAAGAACATCCCTTACTCTCTG 1022
||||| 1179 GACTGCCCTGTGATGCAAGGCGAGTACTTCGCTCTCCCTGGCCCACTACCACTAGC 1238
||||| 1023 GGCAGCTTAGCTGCGTGAAGGCCACCACTACGCGCCCTGGCCCACTACTTCACTG 1082
||||| 1239 CATGGCCCTCTGCAAGGCTCCCGAGCGACCGAGGAGAGTCCCGCAGCAGCAGAGGT 1298
||||| 1083 CATCTCTCTATCGACCACCAAGGTGAAGCCAGGACCGGATCTGGACCAACCAAGGAGAGT 1142
||||| 1299 CTTCTCTGAGC-----CCCCACCTCTTAAGCCCGGAGGCCCTGTGC--TGCCGAGGA 1352
||||| 1143 CCGTCCAGCTCTACGACCACATGCGAGAGGGGCTGACACCTTGGCCCACTGSAAGAA 1202
||||| 1353 GCTGGAGAGCGCAGGAGCTTGGCAAGGCACACTGAAGCGTGCATCTCTGGGCGAGGA 1412
||||| 1203 TGATCAGCAGCGCCGACAGCTGGGAGTCCCACTTGGCAGAGCCATGGCTCATCAGGA 1262
||||| 1413 GGAGCGCTGCGGTGACGCGCTGTGCGCGTCTGCGGAGGTGAGCACTGCTTGGGCG 1472
||||| 1263 GGAGTCCGTGCGGAGGCGAGCTCTGCAAGAAGCTGCGGAGCATTTAGGTGCTACAGAA 1322
||||| 1473 TGTGATCTCCAGAGCGCTGAGCGCTCACTGCGCAAGTATCGGAGCTCGACCGTGAGGA 1532
||||| 1323 GGTGCTGTGCGCGCAGGAAACGCTCCCGGTACGTCACGCCCAACCAAGGAGGAGGA 1382
||||| 1533 TGACTTCTGTGAGGCTGCGGAGGCCCGCGGACATCCAGCTTAAGACCCCAAGAGCCAGA 1592
||||| 1383 TGACTGCTGAACCTGTATCGAGCGCCCGCAGTGTGTTGCTTAAACTGAGCAAGAGTTGA 1442
||||| 1593 GGCCAGGATGCCAGCTGTCTCCAGGGAAGGGGCGCTGACATCTTCCATCGCTGGGCGC 1652

Db 1443 CATTATATTGCCCCAGTTTCCAAAGCTGACAGTCAAGGACTTCTTCCAGAAGCTGGGCC 1502
Qy 1553 CTTCTCTGTGTTTCTCAGCCAAAGAACCGGTGCGGCTGTTGGGGCCCGTCCACCTGACCCG 1712
Db 1503 CTTATCTGTGTTTTCGCTTAACAGCGGTGAGCCCTCTCTGAAAGCATCCGCTTCACTGC 1562
Qy 1713 AGGAGAGGGCGGCTTTGSCCTTCAGCTTCGGGGAGACTCGGCTGCTCTCATCGCTGCCGT 1772
Db 1563 AGAAGAAGGGACTTTGGGTTTCACTTGAGAGGGAAGCCGCCCTTCAAGTTCATTCTCT 1622
Qy 1773 CATTCCAGGGAGCCAGCCCGCGGCTGGCTGAGAGAGCGGCACTACATTTGTGTCACT 1832
Db 1623 GATCTCTTACTGCTGCTCGCTCGGTGGCAGGAGCCCGGGAAGAGATTATATTGTCTCCAT 1682
Qy 1833 GAATGGGAGCCATGAGTGGTGGAGACACGCGAGGTGGTGGAGGCTGAGGCTGCG 1892
Db 1683 TCAGCTTTGTGATTTGTAAGTGGCTGACGCTGAGTGAAGTTATGAAGCTGCTGGAAGGTT 1742
Qy 1893 GGGAGAGCGGGCGCCAGCCCTGACAGGTGGTGTCTGCTGCTG 1932
Db 1743 TGGCGAGGACGAGATCGAGATGAAGTCTGTGAGCTCTCTG 1782

RESULT 11

US-10-239-607-51
; Sequence 51, Application US/10239607
; Publication No. US20030219761A1
; GENERAL INFORMATION:
; APPLICANT: Saatchiologlu, Fahri
; TITLE OF INVENTION: No. US20030219761A1el Prostate-Specific or
; TITLE OF INVENTION: Testis-Specific Nucleic Acid Molecules, Polypeptides, and
; TITLE OF INVENTION: Diagnostic and Therapeutic Methods
; FILE REFERENCE: 50218/003002
; CURRENT APPLICATION NUMBER: US/10/239,607
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09410
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,929
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 3526
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-239-607-51

Query Match 22.9%; Score 496.4; DB 16; Length 3526;
Best Local Similarity 56.8%; Pred. No. 1.5e-111;
Matches 977; Conservative 0; Mismatches 731; Indels 12; Gaps 3;

Qy 219 GGGCTGTGACTCCCTGAGCGCAGATCCAGTGGCGCAGCTGCAGAGCCGCGAGGCCAGAT 278
Db 69 GGGCTGTAATCCCTTGCACAAACCGGCCGAGTAATTTGCAGAAATCAAGAGCTGCTTT 128
Qy 279 TCACCAGCAGATTGACAAAGGAGCTGCAGATCGGAGCGGCGCTGAGAACCTTACAGAGC 338
Db 129 GAATCAGCAGATCTGAAAGCCGTGCGATGAGAGCCGAGCGGAAACCTTCTGAAAGT 188
Qy 339 CACAGCAACAAACCGGCTGAGAGAGCGTCCCTGAGAGTGAAGTACGTCACCTCAA 398
Db 189 GGCACAAACTCAAAAGGTGCGGAGCAAGTGCCTGAGCTGAGCTTCGTCACACTCAGA 248
Qy 399 CCTCAGCTGTGAAGAGGAGCTGAGGAGCTCAGCGGTGGCTGAGCCCTGCGCCGCA 458
Db 249 CTTGAGATGCTCAAGGAAGAGCTGAGAGGGCTGAAATCTCCGTTGGCGTCTATCAGAA 308
Qy 459 TGGAGGGAAGCTGTCTACTGTCCCATCATCCCTCGGCTGAGGAGACCAAGAGCT 518
Db 309 CACAGAGGAGGCAATTTACGATTCCTTCTTGGCTGAGGAAACGAAAGAGCT 368
Qy 519 GGACTGTCTACACCGCTGAGGAGCTGATCTCAGTGCATTTTGAGAGAGCGGCGCTC 578

Db 369 CGACTTTGCGTCTCAAGGATTTTATCTCGAACAATTACAGTGAAGATGGCTATT 428
QY 579 CTACGAGGCGAGAAATCAGGAGCTGAGAGCCCTGCGGCGAGGCATGCGGACCCCGAGCG 638
Db 429 ATATGAAGATGAATTCGAGATCTTATGATCTGAGCAAGCTTGTGCGAGCGCTAGCG 488
QY 639 GAATGAGTGGGCGCTGGAGCTGCTACAGCTATTAACAACAGCTGTGTTCTCTGATGC 698
Db 489 GGATGAGGCGGGGTGAATCTGCTGATGACATCTTCAATCCAGCTGGGCTTGTGCGAGAG 548
QY 699 GGGCTTCTTACCGCTGCGAGAGCCTCGGGCTCTTCTTCCATGCTGTAAGCTGCTTAC 758
Db 549 TCGATTTCTTCCGCGCCACAGCGAGATGGAGCTCTCTGTTACCTGATGACTCTCTCAC 608
QY 759 TGGGGTCCCGGCCAGCGCTGCCCTTTCGAGAAAGGCGAGGTTCTCTTCAACAT 818
Db 609 TGGGGTTCGGGTGAGCAGAGAACCTGCTGCTGAGAGAGGCGAGTGTCTGTTCACAC 668
QY 819 CGGTGCTTCCACAGCAGATGGGGCGCGCAGAGCCGCTCTGACCGAGGCTGCCCG 878
Db 669 TGGGGCTCTTACCCAGATTGGAGCCGCGTGCATCGGACAGCAGCGCTGGGCTGGA 728
QY 879 CGCGCTATGAGGCGCTTCCAGAGGCGCTGCGGCTTTCAGCTCTGAGGAGAGCTT 938
Db 729 GAGTGCATAGATGCTTTCAGAGAGCGCGAGGGGTTTAAATTAATCTGAAGACACATT 788
QY 939 CTCCATGCGCGGAGCCAGACATGAGCGCTGCTGCTGCGACTGAGCAGCTCAT 998
Db 789 TACCCATACTCCAAAGTTACGACATGAGCCCTGCCATGCTCAGGCTGTCTGCAAAATGAT 848
QY 999 GATGGCCCGAGGCGCAGGAATGTGTGTGAGGGCTCTTACCACTGCTGCTTCATGGCCCC 1058
Db 849 GCTTGACAAAGCCCAAGAAAGCGTGTGTGAGAAATCAGCTTCTCTG-----GGATCCG 902
QY 1059 CAGAGCTGCTGCGCCAGCTGCGCTGCGGAGGAGCGCGCCAGCTGCGCAGCCAGTA 1118
Db 903 GAATGAATTTCTCATGCTGGTGAAGTGTGCTCAGGAGGCTGTAAAGTGGAGAGGTCTA 962
QY 1119 CAGGCTAGTGCACCGGACCATGCGCCAGCCAGCCCGCTCCAGATAGCTGCTGCTCTGCTG 1178
Db 963 CMAAGCTACAGCGAGCATGAGCCAGCGCGCGGTGAAGAGAACATCCCTTACTCTG 1022
QY 1179 GACTGCGCTGCTGATGTCAAGCGCGAGTACTTCCGCTCCCTGCGCCACTACCACTGAGC 1238
Db 1023 GGCAGCTTAGCTGCTGAGGCGCCACCACTACCGCGGCGCTGCGCCACTTACTTCACTGC 1082
QY 1239 CATGGCGCTTGCGAGCGCTCCCGAGCGAGGAGAGCTCCCGCAGCAGCAGCAGGT 1298
Db 1083 CATCTCTCTCATCGACACACAGGTGAAGCCAGGACCGATCTGGACCAACAGGAGAGTG 1142
QY 1299 CTTCTCTGCGAGC---CCCCACCTCTCTTAAGCCCGCGAGGCGCTGTGC--TGCGCAGGA 1352
Db 1143 CTTGCTCCAGCTCTACGACACATGCGAGAGGCTGACACCTTGGCCACACTGAAGAA 1202
QY 1353 GCTGAGAGGCGGAGGAGCTTGGCAAGGACACCTGAAGCTGCGCATCTCGGGGAGGA 1412
Db 1203 TGATCAGCAGCGCGCAGCTGGGAGAGTCCCACTTGGCAGAGCCATGGCTCATCAGGA 1262
QY 1413 GAGGCGCTGCGGCTGACCGCTGCGCGCTGCTGCGGAGGAGTGGAGCTGCTGCGGC 1472
Db 1263 GAGTGGTGGGAGCGGAGCTCTGCAAGAGCTGCGGAGCATGAGGTGCTTACAGAA 1322
QY 1473 TGTGATCTCCAGACGCTGCGAGCGCTCACTGGCCCAAGTATGCGGAGCTGCGACCTGAGGA 1532
Db 1323 GGTGCTGTGCGCCACAGAGACGCTCCGGCTCAGTACGCCCGCAGCAGCAGGAGGA 1382
QY 1533 TGACTTCTGTGAGGCTGCGAGGCGCGGACATCCAGCTTAAGACCCACAGAGAGCCAGA 1592
Db 1383 TGACTTGTGAACCTGATCGAGCGCCCCAGTGTGTTGCTTAAACTGAGCAAGAGGTGA 1442
QY 1593 GGCAGGATGCGCGCTTCCAGGGGAGGCGGCTGACATCTTCCATCGCTGCGGCC 1652
Db 1443 CATATATGCCCCAGTTCTCAAGCTGACAGTCAAGGAGTCTTCTCCAGAGAGCTGGGCC 1502

QY 1653 CCTGTCTGTGTCTCAGCCAAAGACCGGTGGCGGCTGTGTGGGCGCGCTCCACTGACCCG 1712
Db 1503 CTTATCTGTGTTTTCGGCTTAAACAAGCGGTGAGCGCTCTCCGAAGCATCCGCTTCACTGC 1562
QY 1713 AGGAGAGGCGGCTTTGGCTTCAAGCTTGGGAGAGCTGCGCTGCTCATCGCTGCGCT 1772
Db 1563 AGAAGAAGGGGACTTGGGGTTTCACTTGAGAGGGAACGCCCCGTTTCAAGTTTCACTTCT 1622
QY 1773 CATTCAGGAGCGCAGGCGCGCGCTGCGCTGAGAGGAGGCGACTACATTTGTGTCACT 1832
Db 1623 GATCTTACTGCTCTGCTTGGCTCGGTGCGAGAGCCCGGGAAGGAGATTAATTTGCTCCAT 1682
QY 1833 GAATGGGAGCAGCATGAGGTGTGAGACACGCGGAGAGTGTGACGAGCTGAGGCTGC 1892
Db 1683 TCAGCTTGTGATTTGTAAGTGTGCTGACGTGAGGTATGAGCTGCTGAAGAGCTT 1742
QY 1893 GAGAGGCGGCGCGCGCTGCGCTGAGGTGTGTGCTGCTG 1932
Db 1743 TGGCAGGAGCAGAGATCAGATGAAGTGTGAGCTCTGAGCTCTCCTG 1782

RESULT 12

US-10-092-900A-351

; Sequence 351, Application US/10092900A

; Publication No. US20040043382A1

; GENERAL INFORMATION:

; APPLICANT: Padigaru, Muralidhara

; APPLICANT: Spytek, Kimberly A.

; APPLICANT: Shenoy, Suresh G.

; APPLICANT: Taupier Jr., Raymond J.

; APPLICANT: Pena, Carol E.A.

; APPLICANT: Li, Li

; APPLICANT: Zerhusen, Bryan D.

; APPLICANT: Gusev, Vladimir Y.

; APPLICANT: Ji, Weizhen

; APPLICANT: Gorman, Linda

; APPLICANT: Miller, Charles E.

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Patturajan, Meera

; APPLICANT: Gangolli, Esha A.

; APPLICANT: Vernet, Corine A.M.

; APPLICANT: Guo, Xiaojia Sasha

; APPLICANT: Tchernev,, Velizar T.

; APPLICANT: Fernandes, Elma R.

; APPLICANT: Casman, Stacie J.

; APPLICANT: Malyankar, Uriel M.

; APPLICANT: Gerlach, Valerie

; APPLICANT: Liu, Yi

; APPLICANT: Anderson, David W.

; APPLICANT: Spaderna, Steven K.

; APPLICANT: Carterton, Elina

; APPLICANT: Leite, Mario W.

; APPLICANT: Zhong, Haibong

; APPLICANT: Alsobrook, John P.

; APPLICANT: Lepley, Denise M.

; APPLICANT: Rieger, Daniel K.

; APPLICANT: Burgess, Catherine E.

; TITLE OF INVENTION: No. US20040043382A1e1 Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 21402-290C

; CURRENT APPLICATION NUMBER: US/10/092,900A

; CURRENT FILING DATE: 2002-03-07

; PRIOR APPLICATION NUMBER: USSN 60/274,322

; PRIOR FILING DATE: 2001-03-08

; PRIOR APPLICATION NUMBER: USSN 60/283,675

; PRIOR FILING DATE: 2001-04-13

; PRIOR APPLICATION NUMBER: USSN 60/338,092

; PRIOR FILING DATE: 2001-12-03

; PRIOR APPLICATION NUMBER: USSN 60/274,281

; PRIOR FILING DATE: 2001-03-08

; PRIOR APPLICATION NUMBER: USSN 60/274,191

; PRIOR FILING DATE: 2001-03-08

; PRIOR APPLICATION NUMBER: USSN 60/325,681

APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
FILE REFERENCE: H1-A0106
CURRENT APPLICATION NUMBER: US/10/108,260A
CURRENT FILING DATE: 2002-03-27
NUMBER OF SEQ ID NOS: 5458
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2397
LENGTH: 3500
TYPE: DNA
ORGANISM: Homo sapiens
US-10-108-260A-2397

Query Match 22.8%; Score 494.8; DB 16; Length 3500;
Best Local Similarity 56.7%; Pred. No. 3.7e-111;
Matches 976; Conservative 0; Mismatches 732; Indels 12; Gaps 3;

QY	219	GGGCTGTGACTCCCTGACCGCAGATCCAGTGGCGGCGAGCTGGAGAGCGCGAGGCGCCAGAT	278
DB	105	GGGCTGTAAATCCCTTGCACAAACCGCGCGAGTAAATTTGCAGATCAAAAGAGCTGCTTT	164
QY	279	TCACACAGAGATTGACAAAGAGCTGCAGATGGGACGGCGCTGAGAACTCTACAGAGC	338
DB	165	GAATCAGCAGATCTGAAAGCCATGGATGAGGACCGGAGCGGAACCTTCTGAAGT	224
QY	339	CACGAGCAACACCGGTTGAGAGACGGTCCCTGGAGCTGAGCTACGTCAACTCCAA	398
DB	225	GGCCACAAACTCAAAGTTGGGAGCMAAGTGGCGGTGGAGCTGAGCTTCGTCAACTCAGA	284
QY	399	CTGAGCTGCTGAAGGAGGAGCTGGAGAGCTCAGCGTGGCGTGGACCTTGGCGGCA	458
DB	285	CTGAGATGCTCAAGGAGAGCTGGAGGGCTGAACTCTCGTGGGCGTCTATCAGAA	344
QY	459	TGGAGCGAAGCTGCACTGTCCCATGATCCCTGGGCTTGAAGGAGCAAGAGCT	518
DB	345	CACAGAGGAGGATTAAGATTCCTGATTCCTTGGCTTGAAGGAGCAAGAGCT	404
QY	519	GGACTGGTCTACACCGCTGAAGAGCTGATCAGTGCACTTTGGAGAGGACGGCGCTC	578
DB	405	CGACTTTGCGTGGTCTCAAGGATTTATCTGGAACTTACAGTGAAGATGGCTATT	464
QY	579	CTACAGGAGCAAAATCAGGAGCTGGAGCCCTGGCGAGGCGATCGGACCCCGCG	638
DB	465	ATATGAAGATGAATTCAGATCTTATGGATCTGAGACAAAGTTGTGGAGCCCTAGCG	524
QY	639	GAATCAGTGGCGCTGGAGCTGCTCAGACCTATTACACAGCTGTGCTTCTTGATGC	698
DB	525	GGATGAGGCGGGTGGAACTGCTGATGACATCTTCATCCAGCTGGGCTTTGTGAGAG	584
QY	699	GGCTTCTCTACCCCTGCGAGAGCTCGGGCTCTTTCCATGATGATGATGATGATGAT	758
DB	585	TCGATTTCTCCCGCCACACGCGAGATGGATCTCCTGTTTCACTGGTATGATCTCTCAC	644
QY	759	TGGGTCCCGCCAGCAGCTGCTGCGCTTCGAGAGGCGAGGTTCTCTTCAAT	818
DB	645	CGGGTTCGGTCAAGCAGACAACTGCTGTGGAGAGGCGAGTGTCTCTTCAACAC	704
QY	819	CGGTGCTTCCACACGAGATTGGGCGCGCGCAGGACCGCTCTCGACCGAGGCTGCCCG	878
DB	705	TGGGGCTCTTAACCCAGATTGGGACCCGCTGTGATCGGAGACGCGGCTGGGCTGA	764
QY	879	CCGCGCTATGGAGGCTTCAGAGGCGCGCTGGGCTTTTACGCTCTCGAGGAGAACTT	938
DB	765	GAGTGCCATAGATGCTTTTTCAGAGAGCGCGAGGGTTTTAAATTAATCTGAAAGACAT	824
QY	939	CTCCCATGCGGAGCCGAGACATGAGCGCTGGTCCCTCTGGGCTCTGAGCAGCTCAT	998
DB	825	TACCCATCTCCAAGTTACGACATGAGCCCTGTCATGCTCAGCGTCTCTGCAAAATGAT	884
QY	999	GATGGCCAGGCGCCAGGAATGTGTTTGGAGGCGCTCTCAGCAGCTTCATCGCTGCGCC	1058
DB	885	GCTTGCAAGCCCAAGAAAGCGTGTGAGAAATCAGCCTTCCTG-----GGATCG	938

QY	1059	CAAAGACTGCTGGCCCGCAGCTGGCGCTCGCGCAGAGAGCGCCCGCAGAGTGGGACGAGTA	1118
DB	939	GAATGAATTTCTCATGTGTGTAAGGTGGCTCAGGAGGCTGTAAAGTGGGAGAGTCTTA	998
QY	1119	CAGGCTAGTGACCGGACCATGCGCCAGCCACCTCCACGACTACGTGCTCTCTCTG	1178
DB	999	CAACAGCTACACGAGCCATGAGCCAGGCGCGGTGAAGAGAACATCCCTACTCTCTG	1058
QY	1179	GACTGCCCTGTGTCATGTCAAGCGCGAGTACTTTCGCTCCCTGGCCCACTACAGTAGC	1238
DB	1059	GGCCAGCTTAGCTGCTGAAGGCCACCACTACGCGGCTGGCCCACTACTTCTACATGC	1118
QY	1239	CATGCGCTCTGCGACGGCTCCCGACGACCGAGGAGAGTCCCGACGACGAGAGGT	1298
DB	1119	CATCTCTCTCATCGACCAACCAAGGTGAAGCCAGGACCGGATCTGGACCAACGAGAGAGT	1178
QY	1299	CTTCTGTCAGC-----CCCCCACTCTTAAGCCCGGAGGCGCTGTC--TGCCGACGA	1352
DB	1179	CTGTGCCAGCTCTAGACCAATGACAGAGGGCTGACACCTTTGGCCACACTGAAGAA	1238
QY	1353	GCTGAGGAGCGCAGGCGAGCTTTGGCAAGGCACACCTGAAGCGTGCATCTCTGGGCGAGA	1412
DB	1239	TGATCAGCAGCGCCGACAGCTGGGGAAGTCCCACTTGGCAGAGCATGGGCTCATCAGA	1298
QY	1413	GGAGGCGCTGGGCTGCGACGCTGCGCGCTCTGCGCGAGGTGGAGCTCTCTGGGCG	1472
DB	1299	GGAGTGGTGGGAGGCGAGCTCTGCAAGAAGCTGGGAGCATTGAGGTCTACAGAA	1358
QY	1473	TGTGATCTCCACAGACGCTGCGAGCGCTCACTGCGCAAGTATGCGGAGCTCGACCGTGAGA	1532
DB	1359	GGTGTGTGTGCGCCGACAGGAACGCTCCCGCTACGTACGCGCCAGACACGAGAGAGGA	1418
QY	1533	TGACTTCTGTGAGGCTGCGGAGCGCCCGGACATCCAGCGCTTAAGACCCACAGAACGAGA	1592
DB	1419	TGACCTGTGTAACCTGATCGACGCGCCCGCAGTGTGTGTTGTTAAACTGACCAAGAGTGA	1478
QY	1593	GGCCAGGATGCAACGCTGTCCAGGGGAAGGGGCTGACATCTTCATCGGCTGGGCGC	1652
DB	1479	CATTATATTGCGCCAGTTCTCAAGCTGACAGTCAAGGACTTCTTCCAGAACTGGGCGC	1538
QY	1653	CTGTCTGTGTCTCAGCCAGAACCGGTGGCGGTGGTGGGCGCGCTCCACTGACCCG	1712
DB	1539	CTTATCTGTGTTTTTTCGGCTAACAGCGTGGAGCGCTCTCTCGAAGCATCCGCTTCTAC	1598
QY	1713	AGGAGGCGGCTTTGGCTCACGCTTGGGAGACTGCGCTGCTCTCATCGCTGCGCT	1772
DB	1599	AGAAGAAGGGACTTGGGCTTCACTTGAAGAGGAAACGCGCCCGTTTCAAGTTCACTTCT	1658
QY	1773	CATTCCAGGAGCCAGCGCGCGCTGGCTGAAGAGGAGGCGACTACATTGTGTCTAGT	1832
DB	1659	GGATCTTACTGCTGCTGCTCGGTGGCAGGAGCCCGGAAAGAGATTAATTGTCTCCAT	1718
QY	1833	GAATGGGAGCCATCAGTGTGTGAGACACGCGAGGTGGTGAAGGAGTGAAGCTCTGAAGAGCT	1892
DB	1719	TCAGCTTGTGATTTGATGTTGCTGACGCTGAGTGGGTTATGAAGCTCTGAAGAGCTT	1778
QY	1893	GGGAGAGGCGGCGCGCAGCTGCGAGTGTGTGCTGCTGTG 1932	
DB	1779	TGGCAGGAGCAGATCGAGATGAAGTCTGTGAGCCTCTCTG 1818	

RESULT 14
US-09-895-040A-2
Sequence 2, Application US/09895040A
Patent No. US20020123474A1
GENERAL INFORMATION:
APPLICANT: Shannon, Mark
APPLICANT: Ji, Yonggang
TITLE OF INVENTION: HUMAN GTP-RHO BINDING PROTEIN 2
FILE REFERENCE: AEOMICA-11
CURRENT APPLICATION NUMBER: US/09/895,040A
CURRENT FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: PCT/US01/00666

;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 09/864,761
;; PRIOR FILING DATE: 2001-05-23
;; NUMBER OF SEQ ID NOS: 180
;; SOFTWARE: Aeomica Sequence Listing Engine
;; SEQ ID NO 2
;; LENGTH: 2061
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-895-040A-2

Query Match 22.7%; Score 493.2; DB 9; Length 2061;
Best Local Similarity 56.7%; Pred. No. 8.9e-111;
Matches 975; Conservative 0; Mismatches 733; Indels 12; Gaps 3;
QY 219 GGGCTGTGACTCCCTGACGAGATCCAGTGGCGGAGCTGCGAGCGCGAGGCCAGAT 278
Db 69 GGGCTGTGAAATCCCTTGACAAACCGCGGAGTAAATTCAGAAATCAAGAGCTGCTT 128
QY 279 TCACACAGATTGACAGAGAGCTGAGATCGGACCGGCGCTGAGAACCTTACAGAGC 338
Db 129 GAATCAGCAGATCTGAAAGCGGTGCGGATGAGGATCGGAGCGGAAACCTTCTGAAAGT 188
QY 339 CACCAGCAACAAACCGGTGAGAGAGCGGTGCGCTGAGCTGAGCTAGCTCAACTCAA 398
Db 189 GGCACAAACTCAAGGTGCGGAGCAAGTGGCGCTGAGCTGAGCTTGTCAACTCAGA 248
QY 399 CTGCACTGCTGAAGAGAGAGCTGAGAGAGCTCAGCGGTGGCTGGAACCTTGGCGGCA 458
Db 249 CCTGCAGATGCTCAAGGAAGAGCTGAGAGGCGCTGAACATCTCGGTGGCGGTCTATCAGAA 308
QY 459 TGGGAGCGAGCTGCTCACTGCTCCCATGATCCCTCGGCTGAGAGAGACCGAGCT 518
Db 309 CACAGAGGAGCATTTACGATTCCTCCCTGATTCCTTGGCTGAGGAAACGAAAGACGT 368
QY 519 GGAATGCTTACACCGCTGAGGAGCTGATCTCAGTGCACCTTTGGAGAGACGGCGCTC 578
Db 369 CGACTTTGCACTGCTCTCAAGGATTTATCTTGAACATTAACAGTGAAGTGGCTATTT 428
QY 579 CTAGAGGAGAAATCAGGAGCTGAGAGCGCTTGGCGGAGCGGACATCGGACCCCGGCG 638
Db 429 ATATGAAGATGAATTCAGATCTTATGGATCTGAGACAAAGCTTGTGGAGCGCTAGCGG 488
QY 639 GAATGAGTGGGCTGAGCTGCTCAGCTTATTAACACAGCTGCTGCTTCTGATGC 698
Db 489 GGATGAGCGGGGTGGAATCTGATGATGATCACTTCACTGAGCTGGGCTTTGTCGAGAG 548
QY 699 GCGCTTCTCACCCTGCCAGGAGCTCGGGCTCTTCTTCCATGTTACGACTCGCTTAC 758
Db 549 TCGATTTCTCCGCGCCACACGGCAGATGGGACTCTCTTCACTTGTATGACTCTCTCAC 608
QY 759 TGGGTTCCCGCCAGCAGCTGCTGCGCTTCGAGAGGAGCGCTTCTTCAACAT 818
Db 609 CCGGGTTCCGGTCAAGCAGCAGAACTGCTGCTGAGAGGCGGAGTGTCTGTGTCAACAC 668
QY 819 CGGTGCGCTTCCACACGAGATTGGGGCGCGCAGACCGCTCTCTGCAACCGAGGCTGCCG 878
Db 669 TGGGGCTCTTACACCCAGATTGGGACCCGGTGGATCGGACGCGAGCGCTGGGCTGGA 728

QY 879 CCGCGCTATGAGCGCTTCCAGAGGGCGGCTTGGGGCTTCCAGCCTCTGAGGGAGACTT 938
Db 729 GAGTGCCATAGATGCCCTTCCAGAGCGCGAGGGGTTTAAATTTACCTGAAAGACACTT 788
QY 939 CTCCCATGCGCGAGCGCCAGACATGAGCGCTTCCCTCTGCGCACTTGGAGAGCTCAT 998
Db 789 TACCCATATCCAAAGTTACGACATGAGCCCTGCCATGCTCAGCTGCTCGTCAAAATGAT 848
QY 999 GATGGCCCGAGCCCGAGGAATGTGTTGAGGGCTCTCACCACCTGCTCCATGCGCCCC 1058
Db 849 GCTTGCACAAAGCCCAAGAAACGCTGTTGAGAAAATCAGCCTTCTCTG-----GGATCCG 902
QY 1059 CCAAGACTGCTGCGCCCGAGCTGCGCTGGGCGCAGAGGCGCCGAGGTGGCAGCGAGTA 1118
Db 903 GAATGAATTTCTCATGCTGGTGAAGTGGCTCAGAGGCTCTAAGGTGGAGAGGCTCTA 962
QY 1119 CAGGCTAGTGACCGGACCATGGCCAGCCACCCGCTCCAGACTTACGTGCTTCTCTG 1178
Db 963 CCAACAGCTACACGACGACCATGAGCCAGGCGCGGTGAAGAGAAACATCCCTACTCTG 1022
QY 1179 GACTGCGCTGCTGATGCTCAAGCGCGAGTACTTCCGCTCCCTGGCCCACTACACAGTAGC 1238
Db 1023 GGCAGCTTAGCCTGCGTGAAGGCCACCACTACGCGCGCTGCGCCCACTACTTCACTGC 1082
QY 1239 CATGGCCCTCTGCGAGCGCTCCCGCAGCGACCGAGGAGAGCTCCCGCAGCAGCAGAGGT 1298
Db 1083 CATCTCTCTCATGACCAACAGGTGAAGCAGGACGAGATCTGGACCAACAGGAGAGTG 1142
QY 1299 CTTCTCTGACG-----CCCCCACTCTTAAAGCCCGAGGCGCTGTGTC--TGCGCGAGGA 1352
Db 1143 CCGTCCCGAGCTCTACGACCATGCGCAGAGGGGCTGACACCTTGGCCCACTGAAGAA 1202
QY 1353 GCTGGAGGCGCAGGAGCTTGGCAGGCGACACCTTGAAGGTGCTTCTGGGCGAGGA 1412
Db 1203 TGATCAGCAGCGCGCAGACTGGGGAAGTCCCACTTGCAGAGAGCTTGGCTCATACGA 1262
QY 1413 GAGGCGCTCGCGTGCACGCGCTGTCGCGCTGTCGCGAGGTGACCTGTCTTCCGGC 1472
Db 1263 GAGTGGTGGGAGGCAAGCTCTGCAAGAGCTGCGGAGCATGAGGTGCTACAGAA 1322
QY 1473 TGTGATCTCCAGACGCTGACGCGCTCATGGGCAAGTATGCGAGCTCGACCGTGAAG 1532
Db 1323 GGTGCTGTGCGCGCACAGGAAAGCTCCCGGCTCAGCTAGCGCCAGCACAGGAGGAGGA 1382
QY 1533 TGACTTCTGTGAGGCTCGCGAGGCGCGGACATCCAGCTTAAGCCCAACAGAGCCAGA 1592
Db 1383 TGACCTGCTGAACCTGATCAGCGCCCAAGTGTGTTGTTGTTAAACTGAGCAAGAGTTGA 1442
QY 1593 GGCAGGATGCCAGCGCTGTCCAGGGGAAAGGGCGCTGACATCTTCCATCGGCTGGGCGC 1652
Db 1443 CATTAATTTGCCCGAGTTCTTCAAGCTGACAGTCAAGGACTTCTTCCAGAGCTGGGCGC 1502
QY 1653 CTTGCTGTGTTTCTAGCCAAAGAACCGGTGGCGGCTGGTGGGCGCGCTTCCACTGACCGG 1712
Db 1503 CTTATCTGTGTTTTCGGCTAAACAGGCGGTGAGCGCTCTCTCGAAGCATCGCTTCACTGC 1562
QY 1713 AGGAGGCGCGCTTGGGCTCAGCTTCCGGGAGACTCGCTGCTCTCTCATCGTGGCGT 1772
Db 1563 AGAAGAGGAGGACTTGGGGTTTCACTTGAAGGAAACGCCCCGCTTCAAGTTCACTTCTCT 1622
QY 1773 CATTCAGGAGGAGCAGCGCGCGGCTGGCTGAGAGGAGGCGGAGCTTACATTTGTGCTG 1832
Db 1623 GGATCTTACTGCTCTGCTGCGTGGCAGGAGCGCGGAGAGAGATTTATTTGTCTCCAT 1682
QY 1833 GAATGGGCGCAGCATGAGGTGGTGGAGACAGCGGAGGTGGTGAAGAGCTGAGGCTGC 1892
Db 1683 TCAGCTGTGTGATTTGAAGTGGCTGACGCTGAGTGAAGTTATGAAGCTGCTGAAGAGCTT 1742
QY 1893 GGCAGAGGCGCGCGCGCAGCTGCAAGTGGTGTGCTGCTGCTG 1932
Db 1743 TGGCGAGGACGAGATCGAGATGAAGTGTGTGAGCTCTCTG 1782

```
RESULT 15
; US-09-895-040A-1
; Sequence 1, Application US/09895040A
; Patent No. US20020123474A1
; GENERAL INFORMATION:
; APPLICANT: Shannon, Mark
; APPLICANT: Ji, Yonggang
; TITLE OF INVENTION: HUMAN GTP-RHO BINDING PROTEIN 2
; FILE REFERENCE: AEOMICA-11
; CURRENT APPLICATION NUMBER: US/09/895,040A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 1
; LENGTH: 3484
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-895-040A-1

Query Match      22.7%; Score 493.2; DB 9; Length 3484;
Best Local Similarity 56.7%; Pred. No. 9e-111;
Matches 975; Conservative 0; Mismatches 733; Indels 12; Gaps 3;

QY 219 GGCTCTGACTCCTCAAGCAGATCCAGTGGCGGCCAGCTGCAGAGCGCGCAGAGCCCGCCAGAT 278
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 89 GGCTCTGTAATCCCTTGGCAAAACCGCGCGAGTAAATTGCAGAAATCAAGAGCTGCTTT 148
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 279 TCACACAGCAGATTGACAGAGAGCTGCAGATCGGAGCGGCGTGAGAACCTCTACAGAGC 338
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 149 GAATCAGCAGATCCTGAAACCGCTGGATGAGGATCGGAGCGGAAACCTTCTGAAAGT 208
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 339 CACCAGCAACACCGGCTGAGAGAGACGGTCCGCTCGAGCTGAGCTACGTCACACTCCAA 398
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 209 GGCCACAACCTCAAGGTGGCGAGCAAGTGGGCTGGAGCTGAGCTTCGTCAACTCAGA 268
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 399 CTGCGAGCTGCTGAAGAGAGCTGAGAGAGCTCAGCGGTGGCGTGAGACCTCGCGCGGCA 458
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 269 CTGCGAGATGCTCAAGAGAGAGCTGAGAGGCGCTGAACATCTCGGTGGCGCTCTATCAGAA 328
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 459 TGGGAGCGAGCTGTCTACTCTCCCATGATCCCGCTGGCGCTGAAGAGACCAAGAGCT 518
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 329 CACAGAGGAGGCAATTTACGATTCCTGATTCCTCTTGGGCTTGAAGGAAACGAAAGAGT 388
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 519 GGAATGCTTACACCGCTGAAGAGAGCTGATCTCAGTGCACTTTGGAGAGGACGCGCCTC 578
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 389 CGACTTTGAGTGGTCTCAAGGATTTATCTCTGGAACATTACAGTGAAGATGCTATTT 448
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 579 CTACGAGCGCAATACAGGAGCTGAGAGGCGCTTCGCGCAGGCCCATGCGACCCCGCCG 638
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 449 ATATGAAGATGAATTCAGATCTTATGGATCTGAGACAAAGCTTGTGCGACGCGCTAGCG 508
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 639 GAATGAGTCGGGCTGAGGCTGCTCAGAGCTATTACAGCCAGCTGTGCTTCTCGGATGC 698
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 509 GGATGAGCGCGGGTGGAACTGTGATGACATCTTATCCAGCTGGGCTTTGTCGAGAG 569
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

699 GGGCTTCTCACCCTCTGCCAGAGCCTCGGGCTCTTCTTCCACTGCTGGTAGCACTCGCTTAC 758
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
569 TCGATTCTTCCCGCCACACAGCAGATGGGACTCTCTGTTCACCTGGTATGACTCTCTCTAC 628
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
759 TGGGGTCCCGGCCAGCAGCGTGGCCCTTGCCAGAAAGGCGCAGCTTCTTCAACAT 818
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
629 CGGGGTTCCGGTTCAGCAGCAGAACCTGTGCTGGAGAAAGCCAGTGCTCTTCAACAC 688
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
819 CGGTGCTCCACACGACAGATTGGGGCGCGCAGAGCCGCTCTCGACCGAGGGTGC CG 878
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
689 TGGGGCCCTCTACACCCAGATTGGGACCCGGTGCATCGGACAGACAGAGCTGGGCTGGA 748
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
879 CGCGCTATGAGAGGCTTCCAGAGGGCGCTGGGGCTTTCAGCTTCCAGAGGAGACTT 938
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
749 GAGTGCCATAGATGCCCTTTCAGAGAGCGCGAGGGGTTTAAATACCTGAAAGACACAT 808
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
939 CTCCCATGCGCGAGCGCCAGACATGAGCGCTCGCTCCCTCTCGCAGCTGGAGCAGCTCAT 998
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
809 TACCCATCTCCAAGTTACGACATGAGCCCTGCCATGCTCAGGGTCTCGTCAATATGAT 868
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
999 GATGCGCCAGGCGCCAGGAAATGTGTGTTGAGGGCTCTCACCACTGCTCCATGGCC 1058
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
869 GCTTGACACAGCCCAAGAAAGCGTGTTCAGAAATCAGCCTTCCTG-----GGATCG 922
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1059 CCAAGACTGCTGGCCAGCTGGCGCTGGCGCAGAGAGGCGCCAGCTGGCAGCCAGTA 1118
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
923 GAATGAATTTCTTCATGCTGTGAAGTGGCTCAGGAGGCTGTAAAGGTGGAGAGTCTA 982
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1119 CAGGCTAGTGACCGGACCATGCGCCAGCCACCGCTCCAGCAGTACGTGCTCTCTCTG 1178
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
983 CCAACAGCTACACGAGCCATGAGCGCGCGGTGAAGAGACATCCCTTACTCTCTG 1042
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1179 GACTGCCCTGTGTCATGTCAAGCGCAGTACTTCCGCTCCCTGGGCCCATACCACTAGC 1238
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1043 GGCAGCTTAGCTGTGCTGAAGCGCCACCACTACGGCGGCTTGGCCCACTTCTTCACTGC 1102
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1239 CATGGCCCTCTGGGAGGCTCCCCAGCGACCGAGGAGAGCTCCCCACCGCAGCAGGT 1298
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1103 CATCTCTCTCATGACACCAGGTGAAGCCAGGACAGCATCTGGACACCGAGGAGAGTG 1162
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1299 CTTCTGTGAGC---CCCCACCTCTCTAAAGCCCGAGGCCCTGTGC--TGCGCAGGA 1352
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1163 CTGTCCCAGCTCTACACACCATGCGAGGGCTGACACCTTGGCCACACTGAGAA 1222
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1353 GCTGGAGGAGCGCAGCAGCTTGGCAAGGACACACTGAAGGCTGCATCTCTGGGGCAGGA 1412
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1223 TGATCAGAGCGCGCAGCAGCTGGGGAAGTCCCACTTGGCGCAGAGCCATGCTCATCAGA 1282
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1413 GAGGGCTGCGGCTGCA CGCCCTGTGCGCGCTCTGCGGAGGTGGAGCTGCTTGGGC 1472
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1283 GAGTGGTGGGAGGCAAGCCTCTGCAAGAAAGCTGGGAGCATTTGAGGTGCTTACAGAA 1342
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1473 TGTGATCTCCAGACCTGCGAGCGCTCACTGGCGAAAGTATGCGAGCTCGACCGTGAGGA 1532
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1343 GGTGCTGTGCGCGCAGAGAGCTCCCGGCTCAGCTAGCCCGCCAGCAGGAGGAGGA 1402
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1533 TGACTTCTGTGAGGCTGCGAGGCGCCCGGACATCCAGCTTAAAGCCACAGAGCCAGA 1592
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1403 TGACCTGTGAACTGTGATGACGCGCCCGCTGTGTGTGTCTTAAACTGAGCAGAGGTTGA 1462
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1593 GGCAGATGCGCAGGCTGTCCAGGGGAGGGGCTGACATCTTCCATCGCTGGGGCC 1652
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1463 CATTATATGGCCCACTTCTCAAGTGAAGTCAAGTCAAGGACTTCTTCCAGAGCTGGGCC 1522
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1653 CTTGTCTGTCTCAGCGCAAGAACCGGTGGCGGCTGGTGGGCGCCGTCCACCTGACCCG 1712
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1523 CTTATCTGTGTTTTGCGCTAACAGCGGTGGAAGCTCTCTCTCTGAGGATCGCTTCACTGC 1582
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1713 AGGAGAGGGCGGCTTTGGCTCAGCTCGGGGAGACTCGCTCTGTCTCTCATCGCTGCCGT 1772
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1583 AGAAGAGGGGACTTGGGGTTCACTTGAGAGGGAAGCGCCCGCTTCAAGTTCACTTCTCT 1642
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1773 CATTCCAGGAGCCAGCGCCCGGCGCTGSCCTGAAGAGGGGCGGACTACATTGTGTCAGT 1832
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db	1643	GGATCCTTACTGCTCTGCTCGCTCGGTGGCAGGAGCCCGGGAAGGAGATTATATTGTCTCCAT	1702
Qy	1833	GAATGGGAGCCCATGCAGGTGGTGGAGACACGGGAGGTGGTGCACGGAGCTGAAGGCTGC	1892
Db	1703	TCAGCTTGTGGATTGTAAAGTGGCTGACGCTGAGTGAGGTTATGAAGCTGCTGAAGAGCTT	1762
Qy	1893	GGGAGAGCGGGCGCCAGCTGCAGGTGGTGTGCTGCTG	1932
Db	1763	TGGCGAGGACGAGATCGAGATGAAGTCGTGAGCCTCCTG	1802

Search completed: July 13, 2004, 12:02:44
Job time : 887.939 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2004, 20:25:00 ; Search time 3670.18 Seconds
(without alignments)
17647.909 Million cell updates/sec

Title: US-10-697-266-1_COPY_103_2271

Perfect score: 2169

Sequence: 1 agatcctggagagagggcc.....tgaagcaccaggggtggccg 2169

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hcc:*

9: gb_est1:*

10: gb_est2:*

11: gb_hcc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_man:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pmg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	833.2	38.4	1070	13	BX400896
2	802.4	37.0	1001	13	BQ072325
3	796.4	36.7	1135	11	BC012135
4	787.6	36.3	899	13	BUI71128

c 5	761	35.1	1017	13	BX444630
c 6	738.6	34.1	1201	13	BX447754
c 7	722	33.3	1127	13	BX341691
8	716.6	33.0	1201	13	BX375208
9	715.8	33.0	1201	13	BX405203
c 10	708.6	32.7	1201	9	AL578313
11	624.4	28.8	1201	13	BX384556
12	613.2	28.3	1243	12	BM928275
13	605.6	27.9	679	10	BE251330
c 14	571	26.3	775	14	CA432432
15	556	25.6	1067	12	BG339711
16	544.4	25.1	3058	11	AK004849
17	541.2	25.0	3351	11	AK050214
18	533.6	24.6	830	12	BI518788
19	526	24.3	986	12	BI770906
20	515.6	23.8	1109	12	BM548005
21	513.2	23.7	793	12	BI103148
22	506	23.3	1201	9	AL555948
23	502.4	23.2	926	13	BQ880416
c 24	499.4	23.0	703	14	CB851722
c 25	495.4	22.8	707	14	CB850862
c 26	468.8	21.6	493	12	BM193145
c 27	468.6	21.6	709	10	BE408252
28	467.8	21.6	609	14	CB216237
c 29	464.4	21.4	494	12	BM153139
30	461	21.3	755	13	BQ179567
c 31	459.8	21.2	650	13	BU629871
c 32	458.6	21.1	810	10	BF607519
33	446	20.6	1408	12	BM475275
34	426	19.6	1021	13	BX444631
35	423.2	19.5	721	13	BQ442874
36	416	19.2	607	14	CD513476
37	411	18.9	955	13	BQ927119
c 38	404.8	18.7	612	9	AI917311
39	403	18.6	999	12	BM802482
40	401	18.5	875	13	BQ944101
41	386.8	17.8	793	13	BX873217
42	385.2	17.8	671	12	BM743041
c 43	366	16.9	611	10	BE328691
44	356.8	16.4	539	12	BM839050
c 45	345.2	15.9	542	10	AW874410

ALIGNMENTS

RESULT 1	BX400896/c	BX400896	Homo sapiens HELA CELLS	1070 bp	mRNA	linear	EST 13-MAY-2003
LOCUS	BX400896	cdna clone CS0DK004YE05	3-PRIME, mRNA sequence.				
DEFINITION	ACCESION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
	BX400896	BX400896.1	GI:30630225	EST.	Homo sapiens (human)	1 (bases 1 to 1070)	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
							Unpublished (2001)
							Full-length cDNA libraries and normalization
							Contact: Genoscope
							Genoscope - Centre National de Sequencage
							BP 191 91006 EVRY cedex - France
							Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
							Library was constructed by Life Technologies, a division of
							Invitrogen. This sequence belongs to sequence cluster 4751.f For
							more information about this cluster, see
							http://www.genoscope.cns.fr/
							cgi-bin/cluster.cgi?seq=CS0DK004AC03NP1&cluster=4751.f. Contact :
							Feng Liang Email : fliang@lifetech.com URL :
							http://fulllength.invitrogen.com/ InVivoGen Corporation 1600

FEATURES		Faraday Avenue Genoscope sequence ID : CS0DK004AC03NP1.	
source	Location/Qualifiers	1..1070	
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CS0DK004YR05"		
ORIGIN	/cell_type="HELA"		
	/cell_lines="HELA"		
	/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"		
	/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."		
Query Match		38.4%;	Score 833.2; DB 13; Length 1070;
Best Local Similarity		97.9%;	Pred. No. 5.6e-134;
Matches		843; Conservative	9; Mismatches 8; Indels 1; Gaps 1;
QY	1309	CCCCCAGCTCTCTAAGCCCGAGGCTGTGTCGCCGAGAGCTGAGGAGCGCAGG	1368
DB	974	SCCCACACCTCTTAACCCGAGCCCTGT-YTCCCCAAGAGCTGAGGAGGCGCAGS	916
QY	1369	CAGCTTGGCAAGCACACTGAAGCGTGCATCCTCTGGGCGAGGAGCGCTGCGGCTG	1428
DB	915	CASCTTGGCAAGACACACTGAAGCGTGCATCCTCTGGGCGAGGAGCGCTGCGGCTG	856
QY	1429	CAGCGCCTGTGCGGCTCTCGCGGAGGTGACCTGCTTCGGGCTGTGATCTCCAGAGC	1488
DB	855	CAGCGCCTGTGCGGCTCTCGCGGAGGTGACCTGCTTCGGGCTGTGATCTCCAGAGC	796
QY	1489	CTGCAAGCGCTACTGGCCAAAGTATGCGGAGCTCGACCGTGAGATGACTTCTGTAGGCT	1548
DB	795	CTGCAAGCGCTACTGGCCAAAGTATGCGGAGCTCGACCGTGAGATGACTTCTGTAGGCT	736
QY	1549	GGCGAGGCCCCGAGATCCAGCTTAAGCCACCCAGAGCCAGAGCCAGGATGCCAGC	1608
DB	735	GGCGAGGCCCCGAGATCCAGCTTAAGCCACCCAGAGCCAGAGCCAGGATGCCAGC	676
QY	1609	CTGTCCAGGGAAGGGGCTGACATCTTCCATCGGCTGGGGCCCCGTGCTGTCTTCA	1668
DB	675	CTGTCCAGGGAAGGGGCTGACATCTTCCATCGGCTGGGGCCCCGTGCTGTCTTCA	616
QY	1669	GCAAGAACGGTGGCGGTGTGTGGGCCCCGTCCACCTGACCGGAGGAGGCGGCTTT	1728
DB	615	GCAAGAACGGTGGCGGTGTGTGGGCCCCGTCCACCTGACCGGAGGAGGCGGCTTT	556
QY	1729	GGCCTCAGCTTTCGGGAGACTCGCTGTCTCATCGCTGCCGTCAATTCAGGAGCCAG	1788
DB	555	GGCCTCAGCTTTCGGGAGACTCGCTGTCTCATCGCTGCCGTCAATTCAGGAGCCAG	496
QY	1789	GCCGCGCGGCTGGCTGTAAGGAGGCGACTACATTGTCTGATGTAATGGGAGCCATGC	1848
DB	495	GCCGCGCGGCTGGCTGTAAGGAGGCGACTACATTGTCTGATGTAATGGGAGCCATGC	436
QY	1849	AGGTGTGAGACACCGGAGGTGTGACGAGCTGAAGGCTCGGGAGAGCGGCGCC	1908
DB	435	AGGTGTGAGACACCGGAGGTGTGACGAGCTGAAGGCTCGGGAGAGCGGCGCC	376
QY	1909	AGCTGACGTTGTGTGTGCTGCCAGCTCTAGACTGCCAGCTTGGGGACCGCCGG	1968
DB	375	AGCTGACGTTGTGTGTGCTGCCAGCTCTAGACTGCCAGCTTGGGGACCGCCGG	316
QY	1969	CCCGTCTCTGGGCCCCAAGGGGCTTCTAAGGAGCCAGAGGAGCATGTTTCAAGACC	2028
DB	315	CCCGTCTCTGGGCCCCAAGGGGCTTCTAAGGAGCCAGAGGAGCATGTTTCAAGACC	256
QY	2029	CCGGATTCACGTGGGCGAGTCCCGGCCCTCTCACTGGAGCCGAAGGCCAGCAG	2088
DB	255	CCGGATTCACGTGGGCGAGTCCCGGCCCTCTCTTAACTGGAGCCGAAGGCCAGCAG	196
QY	2089	GGCAAGACTGGAGGCTGCCCCAGCCCTGTGTGCCCCAGTGAAGCCAGCTCCGCCCTCATCC	2148
Db		195	GGCAGACTGGAGGCTGCCCCAGCCCTGTGTGCCCCAGTGAAGCCAGCTCCGCCCTCATCC
QY		2149	TTGAAGCACCCAGGCTGGCCG 2169
Db		135	TTGAAGCACCCAGGCTGGCCG 115
RESULT 2			
BQ072325		1001 bp mRNA linear EST 02-APR-2002	
LOCUS		AGENCOURT 6838996 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5761460	
DEFINITION		5' mRNA Sequence.	
ACCESSION		BQ072325	
VERSION		BQ072325.1 GI:19901371	
KEYWORDS		EST.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE		1 (bases 1 to 1001)	
AUTHORS		NIH-MGC http://mgc.hci.nih.gov/.	
TITLE		National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL		Unpublished (1999)	
COMMENT		Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM12810 row: b column: 21 High quality sequence stop: 623.	
FEATURES		Location/Qualifiers	
source		1..1001	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:5761460"	
		/lab_host="DH10B"	
		/clone_lib="NIH_MGC_122"	
		/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 Kb, insert size range 1-3 Kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH_MGC Library."	
ORIGIN			
Query Match		37.0%;	Score 802.4; DB 13; Length 1001;
Best Local Similarity		96.4%;	Pred. No. 1.2e-128;
Matches		864; Conservative	0; Mismatches 26; Indels 6; Gaps 4;
QY	219	GGGCTGTGACTCCCTGAOCAGATCCAGTGGCGCCAGCTGCAGAGCCGAGGCCCCAGAT	278
DB	42	GGGCTGTGACTCCCTGACGACAGATCCAGTGGCGCCAGCTGCAGAGCCGAGGCCCCAGAT	101
QY	279	TCACGACGAGATTGCAAGGAGCTGCAGATCGGCGCTGAGAACTCTTACAGAGC	338
DB	102	TCACGACGAGATTGCAAGGAGCTGCAGATCGGCGCTGAGAACTCTTACAGAGC	161
QY	339	CACCGACAAACCGGGTGAAGAGACGCGTCCGCTGGAGCTGAGTACCTCACTCCAA	398
DB	162	CACCGACAAACCGGGTGAAGAGACGCGTCCGCTGGAGCTGAGTACCTCACTCCAA	221
QY	399	CTGCAAGCTGTGAAGGAGGAGCTGAGGAGCTCAGCGTGGCTGGACCTGCGCGCA	458

```

222 CCTGCAGCTGCTGAAGAGGAGCTGGAGAGCTACGCGGTGGGACCTTGGCGCGCA 281
Qy TGGGAGCGAAAGCTGTCACTGTCCCATGATCCCTCGGCGCTGAAGGAGACCAAGGAGCT 518
Db TGGGAGCGAAAGCTGTCACTGTCCCATGATCCCTCGGCGCTGAAGGAGACCAAGGAGCT 341
Qy GGACTGCTGTACAGCGCTGAAGAGAGCTGATCTCAGTGTCACTTTGAGAGGACGGCGCTC 578
Db GGACTGCTGTACAGCGCTGAAGAGAGCTGATCTCAGTGTCACTTTGAGAGGACGGCGCTC 401
Qy CTACGAGCGCAAAATCAGGAGCTGGAGCGCTCGGCGAGGCCATGCGGACCCCGAGCG 638
Db CTACGAGCGCAAAATCAGGAGCTGGAGCGCTCGGCGAGGCCATGCGGACCCCGAGCG 461
Qy GAATGAGTGGCGCTGGAGCTGTCTACAGCGCTATTACAACAGAGCTGTCTTCTTGGATGC 698
Db GAATGAGTGGCGCTGGAGCTGTCTACAGCGCTATTACAACAGAGCTGTCTTCTTGGATGC 521
Qy GCGCTTCTCACCCTGCCAGGAGCTCGGCGCTTCTTCCACTGGTACGACTCGCTTAC 758
Db GCGCTTCTCACCCTGCCAGGAGCTCGGCGCTTCTTCCACTGGTACGACTCGCTTAC 581
Qy TGGGGTCCCGGCCAGCAGCGTGCCTTGGCTTCGAGAGGCGAGCTTCTTCAACAT 818
Db TGGGGTCCCGGCCAGCAGCGTGCCTTGGCTTCGAGAGGCGAGCTTCTTCAACAT 641
Qy CGGTGCTTCCACACGAGATTTGGGGCGCGCAGGACCGCTTCTTCCACCGAGGCGCGCG 878
Db CGGTGCTTCCACACGAGATTTGGGGCGCGCAGGACCGCTTCTTCCACCGAGGCGCGCG 701
Qy CCGCGCTATGAGCGCTTTCAGAGGCGCGCTTGGGGCGCGCAGGACCGCTTCTTCAACAT 937
Db CCGCGCTATGAGCGCTTTCAGAGGCGCGCTTGGGGCGCGCAGGACCGCTTCTTCAACAT 761
Qy TCTCCCATGCGCGC-AGCCAGACATGAGCGCTCGCTTCTTCCAGGAGGCGCGCGAGCGTC 996
Db TCTCCCATGCGCGC-AGCCAGACATGAGCGCTCGCTTCTTCCAGGAGGCGCGCGAGCGTC 821
Qy ATGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1054
Db ATGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 881
Qy CCCCCCAGAGCTGCTGG--CCAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1108
Db CCCCCCAGAGCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 937

RESULT 3
BC012135
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BC012135.1 GI:15082433
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1135)
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,
Scheetz T.E., Brownstein M.J., Urdin T.B., Toshiyuki S.,
Carninci P., Prange C., Raja S.S., Loquellano N.A., Peters G.J.,
Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J.,
McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,
Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettman M., Madan A., Rodriguez S.,
```

```

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E.,
Butcher, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477332
2 (bases 1 to 1135)
Strausberg, R.
Direct Submission
Submitted (02-AUG-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcaps-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline
Schein, Duane Smallos, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
Series: IRAL Plate: 29 Row: d Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 19882236
This clone has the following problem: retained intron.
```

FEATURES

```

Location/Qualifiers
1..1135
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4555079"
/tissue_type="Uterus, leiomyosarcoma"
/clone_lib="NIH MGC 46"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
```

ORIGIN

```

Query Match      36.7%   Score 796.4;   DB 11;   Length 1135;
Best Local Similarity 99.9%;   Pred. No. 1.3e-127;
Matches 797;   Conservative 0;   Mismatches 1;   Indels 0;   Gaps 0;

Qy 219 GGGCTGTGACTCCCTGACGAGAGCTGCGAGCGCGCTGAGAACTCTACAGAGC 278
Db 162 GGGCTGTGACTCCCTGACGAGATTCAGTGGCGCCAGCTGCAGAGCGCGCCAGAT 221
Qy 279 TCACGAGCAGATTGACAGAGCTGCGAGATGCGAGCGCGCTGAGAACTCTACAGAGC 338
Db 222 TCACGAGCAGATTGACAGAGCTGCGAGATGCGAGCGCGCGCTGAGAACTCTACAGAGC 281
Qy 339 CACGAGCAACACCGGGTGAAGAGAGCGGTGCGCTTGGAGCTGAGTACGCTCAACTCCAA 398
Db 282 CACGAGCAACACCGGGTGAAGAGAGCGGTGCGCTTGGAGCTGAGTACGCTCAACTCCAA 341
Qy 399 CTTGAGCTGTGAGAGGAGAGCTGAGAGAGCTACGCGTGGCGTGGAGACCTGGCGCGCA 458
Db 342 CTTGAGCTGTGAGAGGAGAGCTGAGAGAGCTACGCGTGGCGTGGAGACCTGGCGCGCA 401
```

QY	459	TGGGAGCGAAGCTGTCACTGTCTCCCATGATCCCTTGGCCCTGAGGAGACCAAGAGCT	518
Db	402	TGGGAGCGAAGCTGTCACTGTCTCCCATGATCCCTTGGCCCTGAGGAGACCAAGAGCT	461
QY	519	GGACTGGTCTACACCGCTGAAGAGCTGATCTCAGTGCACCTTTGGAGAGGACGGCGCTTC	578
Db	462	GGACTGGTCTACACCGCTGAAGAGCTGATCTCAGTGCACCTTTGGAGAGGACGGCGCTTC	521
QY	579	CTACGAGGCGAGAAATCAGGAGCTGGAGCCCTGGCGCAGGCGCATGGGACCCCA3CG	638
Db	522	CTACGAGGCGAGAAATCAGGAGCTGGAGCCCTGGCGCAGGCGCATGGGACCCCA3CCA	581
QY	639	GAATGAGTGGGCTTGGAGCTGCTCACAGCTATTACACACGCTGTGCTTCTTGGATGC	698
Db	582	GAATGAGTGGGCTTGGAGCTGCTCACAGCTATTACACACGCTGTGCTTCTTGGATGC	641
QY	699	GGGCTTCTTCCATCCCTGCCAGGAGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC	758
Db	642	GGGCTTCTTCCATCCCTGCCAGGAGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC	701
QY	759	TGGGFTCCCGCCCGCAGAGCTGGCCCTTGGCCCTTGCAGAAAGGCGAGGTTCTCTTCAACAT	818
Db	702	TGGGFTCCCGCCCGCAGAGCTGGCCCTTGGCCCTTGCAGAAAGGCGAGGTTCTCTTCAACAT	761
QY	819	CGTGCCCTTCCACACGACATTTGGGGCGCGCAGGACCGCTCTGACACGAGGCTCCCG	878
Db	762	CGTGCCCTTCCACACGACATTTGGGGCGCGCAGGACCGCTCTGACACGAGGCTCCCG	821
QY	879	CGCGCTATGGAGGCTTCCAGAGGCGCTGGGGCTTTCAGCTCTCTTGGAGGAGACTT	938
Db	822	CGCGCTATGGAGGCTTCCAGAGGCGCTGGGGCTTTCAGCTCTCTTGGAGGAGACTT	881
QY	939	CTCCATGCGCGAGCCCGCAGACATGAGCGTGGCTCCCTCTGGGCTCTGAGGAGTCAAT	998
Db	882	CTCCATGCGCGAGCCCGCAGACATGAGCGTGGCTCCCTCTGGGCTCTGAGGAGTCAAT	941
QY	999	GATGGCCCGAGGCCAGGA 1016	
Db	942	GATGGCCCGAGGCCAGGA 959	
RESULT 4			
LOCUS	BUI71128	899 bp mRNA linear EST 04-SEP-2002	
DEFINITION	AGENCOURT_7935660 NIH_MGC 67 Homo sapiens cDNA clone IMAGE:6146655		
ACCESSION	BUI71128	5', mRNA sequence.	
VERSION	BUI71128.1	GI:22685112	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1	(bases 1 to 899)	
AUTHORS	NIH-MGC	http://mgi.nci.nih.gov/	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L1AM13475 row: d column: 16 High quality sequence stop: 563. Location/Qualifiers 1. .899 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606"		
FEATURES	source		

/clone="IMAGE:6146655"			
/tissue type="retinoblastoma"			
/lab host="PDH10B (phage-resistant)"			
/clone_lib="NIH_MGC_87"			
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.75 kb. Library constructed by Life Technologies."			
ORIGIN			
Query Match 36.3%; Score 787.6; DB 13; Length 899;			
Best Local Similarity 98.6%; Pred. No. 4e-126;			
Matches 826; Conservative 0; Mismatches 9; Indels 3; Gaps 3;			
QY	1023	GTTCAGGCGCTCTCACACCTGCTCCATGCGCCGCCCAAGACTGCTTGCCGACAGTGGC	1082
Db	1	GTTCAGGCGCTCTCACACCTGCTCCATGCGCCGCCCAAGACTGCTTGCCGACAGTGGC	60
QY	1083	CCTGCGCAGGAGGCGCCAGGTGGCAGCCGAGTACAGGCTAGTGACACCGACCATGGC	1142
Db	61	CCTGCGCAGGAGGCGCCAGGTGGCAGCCGAGTACAGGCTAGTGACACCGACCATGGC	120
QY	1143	CAAGCACCCTGTCAGACTAGTGGCTGTCTTGGACTGCTGTCGTCGTCGTCGTCGTC	1202
Db	121	CAAGCACCCTGTCAGACTAGTGGCTGTCTTGGACTGCTGTCGTCGTCGTCGTCGTC	180
QY	1203	CGAGTACTTCCCTCCCTGGCCCACTACCACTGAGCCATGGCCCTCTGCGACGGCTCC	1262
Db	181	CGAGTACTTCCCTCCCTGGCCCACTACCACTGAGCCATGGCCCTCTGCGACGGCTCC	240
QY	1263	AGCGACCGAGGAGAGCTCCCAACGACGAGCAGGCTTCTTGGACGCCCTCCACCTCT	1322
Db	241	AGCGACCGAGGAGAGCTCCCAACGACGAGCAGGCTTCTTGGACGCCCTCCACCTCT	300
QY	1323	TAAGCCCGAGGCGCTGTGCTGCGCAGAGCTGGAGAGCGCAGCAGCTTGGCAGGC	1382
Db	301	TAAGCCCGAGGCGCTGTGCTGCGCAGAGCTGGAGAGCGCAGCAGCTTGGCAGGC	360
QY	1383	ACACCTGGAAGCTGCCATCTGCGGGCAGGAGGCGCTGCGGCTGCAGCGCTGTGCCG	1442
Db	361	ACACCTGGAAGCTGCCATCTGCGGGCAGGAGGCGCTGCGGCTGCAGCGCTGTGCCG	420
QY	1443	CCTCTCTGCGCAGGTGGACCTGCTTCGGCTGTGATCTCCAGACGCTGACGCGCTCACT	1502
Db	421	CCTCTCTGCGCAGGTGGACCTGCTTCGGCTGTGATCTCCAGACGCTGACGCGCTCACT	480
QY	1503	GGCAGAGTATGGAGCTGACCGTGAGTGAATCTTCTGTGAGGCTGCGAGGCCCGGA	1562
Db	481	GGCAGAGTATGGAGCTGACCGTGAGTGAATCTTCTGTGAGGCTGCGAGGCCCGGA	540
QY	1563	CATCCAGCCTTAAGACCCACAGAACGAGGCGCAGGATGCCACGCTGTCCAGGGAA	1622
Db	541	CATCCAGCCTTAAGACCCACAGAACGAGGCGCAGGATGCCACGCTGTCCAGGGAA	600
QY	1623	GGGG-CCTGACATCTTCCATPCGCTGGGCGCTGTGTCGTGTCCTCAGCCAGAACCGGT	1681
Db	601	GGGG-CCTGACATCTTCCATPCGCTGGGCGCTGTGTCGTGTCCTCAGCCAGAACCGGT	660
QY	1682	GGCGGCTGTGGGCGCGCTCCACTGACCGCAGGAGGCGGCTTGGCTTCAAGCTTC	1741
Db	661	GGCGGCTGTGGGCGCGCTCCACTGACCGCAGGAGGCGGCTTGGCTTCAAGCTTC	720
QY	1742	GGGGAGACTGCGCTGTCTCATCGCTGCCCTCATCTCCAGGAGCCAGGCCGCGG-CGGCT	1800
Db	721	GGGGAGACTGCGCTGTCTCATCGCTGCCCTCATCTCCAGGAGCCAGGCCGCGGCTT	780
QY	1801	GGCCTGAGAGGAGGCGCAGTACAT-TGTGTCAGTGAATGGCAGCCATGACAGGTGGTGG	1857
Db	781	GGCCTGAGAGGAGGCGCAGTACAT-TGTGTCAGTGAATGGCAGCCATGACAGGTGGTGG	838
RESULT 5			
BX444630/c			

LOCUS	BX444630	1017 bp	mRNA	linear	EST 15-MAY-2003
DEFINITION	BX444630 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone CS0DN004YH13 3-PRIME, mRNA sequence.				
ACCESSION	BX444630				
VERSION	1				
KEYWORDS	EST.				
SOURCE	BX444630.1 GI:30778254				
ORGANISM	Homo sapiens (human)				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1. (bases 1 to 1017)				
TITLE	Li, W.B., Gruber, C., Jessee, J., and Polayes, D.				
COMMENT	Full-length cDNA libraries and normalization Unpublished (2001) Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4751.f For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS0DN004CD07NP1&cluster=4751.f. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DN004CD07NP1. Location/Qualifiers 1. 1017 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DN004YH13" /tissue_type="ADULT BRAIN" /dev_stage="adult" /clone_lib="Homo sapiens ADULT BRAIN" /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."				
FEATURES	source				
ORIGIN	Query Match Best local similarity 35.1%; Score 761; DB 13; Length 1017; Matches 794; Conservative 0; Mismatches 0; Indels 3; Gaps 3;				
QY	1375 GGCAAGGCACACCTGAAGCGCTGCATCCTCGGCGCAGGAGAGCGCTGCGGCTGCACGCC 1434				
Db	938 GGCAAGGCACACCTGAAGCGCTGCATCCTCGGCGCAGGAGAGCGCTGCGGCTGCACGCC 879				
QY	1435 CTGTGCGCGCTCTGCGCGAGGTGGACCTGCTT - CGGGCTGTGATCT - CCCAGACGCTGC 1492				
Db	878 CTGTGCGCGCTCTGCGGAGGTGGACCTGCTTCCGGGCTGTGATCTCCCGACGCGCTGC 819				
QY	1493 AGCGCTCACTGGCCAGTATGCGGAGCTGCACCGTGAGGATGACTTCTGTGAGGCTGCGCG 1552				
Db	818 AGCGCTCACTGGCCAAATATGCGGAGCTGCACCGTGAGGATGACTTCTGTGAGGCTGCGCG 759				
QY	1553 AGGCCCCGACATCCAGCCCTTAAGACCCACAGAGCCAGAGCCAGATGCCAGCGCTGT 1612				
Db	758 AGGCCCCGACATCCAGCCCTTAAGACCCACAGAGCCAGAGCCAGATGCCAGCGCTGT 699				
QY	1613 CCCAGGGGAAGGGCGCTGACATCTTCATCGGCTGGGGCCCTGTCTGTGTTCTCAGCCA 1672				
Db	698 CCCAGGGGAAGGGCGCTGACATCTTCATCGGCTGGGGCCCTGTCTGTGTTCTCAGCCA 639				
QY	1673 AGAACCGGTGGCGGCTGGTGGGGCCCGTCCACTGACCCGAGGAGAGGGCGCTTTGGCC 1732				
Db	638 AGAACCGGTGGCGGCTGGTGGGGCCCGTCCACTGACCCGAGGAGAGGGCGCTTTGGCC 579				
QY	1733 TCACGCTTCGGGAGACTCGCTGTCTCTATCTCGCTGCCGTCAATCCAGGAGCCAGGCGG 1792				
Db	578 TCACGCTTCGGGAGACTCGCTGTCTCTATCTCGCTGCCGTCAATCCAGGAGCCAGGCGG 519				

Db	805	TGGGGT-CCGGCCACGACGCTGGCCCTTCGAGAAGGCGCAGCGTTCCTCTCAACAT	863
Qy	819	CGGTGCCCTCCACACGACGATTGGGGCGCGCAGACCGCTCTCCACCGAGGGTGCCCG	878
Db	864	CGGTGCCCTCCACACGACGATTGGGGCGCG-CAGACCGCTCTCCACCGAGGT-SCCG	921
Qy	879	CGGGCTATGAGGCGCTTCCAGAGGGCGCGTGGGGCCCTTCAGCCTCTCGAGGAGAACTT	938
Db	922	CCGGCTATGAGG-CTTCCAGAGGGCGCGTGGGGCCCTTCAGCCTCTCGAGGAGAACTT	980
Qy	939	CTCCATGCGCGCGAGCCAGACATGAGCGCTGCGCTCCTCTGCGCATGAGGAGCTCAT	998
Db	981	CTCCATGCGCGCGAGCCAGACATGAGCGCTGCTTCCTTCCTGCGCATGAGGAGCTCAT	1040
Qy	999	GATGGCCAGGCCCGAGGAATGTGTGTTGAGGGCTCTCACCACTGCCTCCATGGCCCC	1058
Db	1041	GATGGCCAGGCCCGAGGAATGTGTTKTKTKKAGGC--TTCACACTGCTGCTGAGCCCC	1098
Qy	1059	CCAA 1062	
Db	1099	CAA 1102	
RESULT 8			
LOCUS	BX375208	1201 bp mRNA linear EST 08-MAY-2003	
DEFINITION	BX375208 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens		
ACCESSION	CDNA clone CS0DC010YG11 5-PRIME, mRNA sequence.		
VERSION	BX375208		
KEYWORDS	BX375208.1 GI:30454376		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4751.f For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS0DC010AD06QPI&cluster=4751.f. Contact : Peng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DC010AD06QPI. Location/Qualifiers 1. .1201 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DC010YG11" /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED" /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED" /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSP6 vector. Library was normalized."		
FEATURES	source		
ORIGIN			
Query Match	33.0%;	Score 716.6;	DB 13; Length 1201;
Best Local Similarity	94.5%;	Pred. No. 9e-114;	
Matches	778; Conservative	16; Mismatches	20; Indels 9; Gaps 5;
Qy	219	GGCTGTGACTCCACGACAGATCCAGTCGCGCCAGCTGCAGAGCCGAGGCCCCAGAT	278
Db	223	GGGCTGTGACTCCCTGACGACAGATCCAGTCGCGCCAGCTGCAGAGCCGAGGCCCCAGAT	282

Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4751.f For more information about this cluster, see

<http://www.genoscope.cns.fr/>

<http://biocluster.cgi?seq=CS0AH001CB05QP1&cluster=4751.f>. Contact : Feng Liang Email : fliang@lifetech.com URL :

<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0AH001CB05QP1.

FEATURES

source

1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0BH001YD09"
/tissue_type="T CELLS (JURKAT CELL LINE)"
/cell_line="JURKAT CELL LINE"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 33.0%; Score 715.8; DB 13; Length 1201;
Best Local Similarity 91.9%; Pred. No. 1.2e-113;
Matches 824; Conservative 6; Mismatches 13; Indels 54; (aps 5;
QY 227 ACTCCCTGACGAGATCCAGTGGCCAGCTGCAGAGCCGAGGGCCAGATTCAACAGC 286
Db 211 ACCCGGGCTGCAGATCCAGTGGCCAGCTGCAGAGCCGCA-GGCCAGATTCAACAGC 269
QY 287 AGATTGACAGAGCTGCAGATCGGACGGCGCTGAGAACCTCTACAGAGCCACAGCA 346
Db 270 AGATTGACAGAGCTGCAGATCGGACGGCGCTGAGAACCTCTACAGAGCCACAGCA 329
QY 347 ACAACGGGTGAGAGAGCGTCCCTGGAGTGAAGTCACTCAACTCAAGCTTCAGC 406
Db 330 ACAACGGGTGAGAGAGCGTCCCTGGAGTGAAGTCACTCAACTCAAGCTTCAGC 389
QY 407 TGCTGAAGGAGAGCTGGAGAGCTCAGCGGTGGCTGGACCTGCGCGCATGGAGCG 466
Db 390 TGCTGAAGGAGAGCTGGAGAGCTCAGCGGTGGCTGGACCTGCGCGCATGGAGCG 449
QY 467 AGCTGTCACTGCCCATGATCCCTGGCTGGCTGAGGAGACACAGGAGCTGGAATGCT 526
Db 450 AA-CTGTCACTGCCCATGATCCCTGGCTGGCTGAGGAGACACAGGAGCTGGAATGCT 508
QY 527 CTACACCGCTGAAGGAGCTGATCTCAGTGCATTTGGAGAGGACGGCGCTCTCTAAGAG 586
Db 509 CTACACCGCTGAAGGAGCTGATCTCAGTGCATTTGGAGAGGACGGCGCTCTCTAAGAG 568
QY 587 CAGAAATCAGGAGCTGGAGGCTCGGGCAGGCCATTCGGACCCCGACCGGAAAGAGT 646
Db 569 CAGAAATCAGGAGCTGGAGGCTCGGGCAGGCCATTCGGACCCCGACCGGAAAGAGT 628
QY 647 CGGGCTGGAGCTGCTCAGAGCTATTACACAGCTGTCTCTCGGAGCGCGTTCC 706
Db 629 CGGGCTGGAGCTGCTCAGAGCTATTACACAGCTGTCTCTCGGAGCGCGTTCC 688
QY 707 TCACCCCTGCAGAGGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTACTGGTCTCC 766
Db 689 TCACCCCTGCAGAGGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTACTGGTCTCC 748
QY 767 CGGCCAGACGCTGCTCGCTTCGAGAGGACGCTTCTTCAACATCGGTCGCC 826
Db 749 CGGCCAGACGCTGCTCGCTTCGAGAGGACGCTTCTTCAACATCGGTCGCC 808
QY 827 TCACACGCGAGATT-GGGGCGCGCAGACCGCTCTCTGACCGAGGCTGCCCGCGCT 885
Db 809 TCACACGCGAGATTGGGGGCGCGCAGACCGCTCTCTGACCGAGGCTGCCCGCGCT 868
QY 886 ATGAGGCGCTTCCAGAGGGCGCTGGGGCTTTCAGCCCTCCTGAGGAGAACTTCTTCCAT 945

Db 869 ATGAGGCGCTTCCAGAGGCGCGCT-----892
QY 946 GGGCCGAGCCAGACATGATGCGTCCCTCTGCGACTGAGCAGCTCATGATGGCC 1005
Db 893 -----GATGAGCGCTGCGTCCCTCTGCGACTGAGCAGCTCATGATGGCC 940
QY 1006 CAGGCCAGGAATGTGTGTTGAGGCGCTCTCACACCTGCTCCATGCGCCCCCAAGAC 1065
Db 941 CAGGCCAGGAATGTGTGTTGAGGCGCTCTCACACCTGCTCCATGCGCCCCCAAGAC 1000
QY 1066 TGCCTGGCCAGCTGCGCTGCGCAGGAGCGCCAGGTCGAGCCCGAGTACAGG 1122
Db 1001 TGCCTGGCCAGCTGCGC---TGGGCGAAGAGCGCCAGTGCAGCCGAGTACAGS 1054
RESULT 10
AL578313/c 1201 bp mRNA linear EST 01-JUN-2003
LOCUS AL578313 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION CDNA clone CS0DK010YM04 3-PRIME, mRNA sequence.
ACCESSION AL578313
VERSION AL578313.2 GI:31316527
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 16, 2001 this sequence version replaced gi:12942264.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4751.f For more information about this cluster, see
<http://www.genoscope.cns.fr/>
<http://biocluster.cgi?seq=CS0DK010BG02NP1&cluster=4751.f>. Contact : Feng Liang Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DK010BG02NP1.
FEATURES
source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK010YM04"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
ORIGIN
Query Match 32.7%; Score 708.6; DB 9; Length 1201;
Best Local Similarity 95.7%; Pred. No. 2.2e-112;
Matches 761; Conservative 8; Mismatches 22; Indels 4; Gaps 4;
QY 1376 GCAAGGCACACCTGAAAGCGTGCATCTCTGGGGCAGGAGGAGGCGCTCGGCTGCACGCC 1435
Db 946 KCTGGCCCCASISGYTTCGCGAGGAGCGCCAGGAGGAGGCGCTC-GCTGCACGCC 888
QY 1436 TGTGCGCGTCTCTGCGAGGTGGAAGTCTTGGGCTGTGATCTCCAGACGCTGAGC 1495
Db 887 TGTGCGCGTCTCC-GCGCAGGTGGAAGTCTTGGGCTGTGATCTCCAGACGCTGAGC 829
QY 1496 GCTCACTGGCCAAAGTATGCGAGCTCGACCGTGAGGATGACTTCTGTGAGGCTGCCGAGG 1555

Db	828	GCTCAGCTGGCCAAAGTATGCGGAGCTGCACCGTGAAGATGACTTCTGTGAAGCTGCCGAGG	769
QY	1556	CCCCGGACATCCAGCCTAAGAACCCACAGAAAGCCAGAGGGCCAGGATGCGACGCTGTGCC	1615
Db	768	CCCCGGACATCCAGCCTAAGAACCCACAGAAAGCCAGAGGGCCAGGATGCGACGCTGTGCC	709
QY	1616	AGGGGAAGGGGCGCTGACATCTTCATCGGCTGGGGCCCTGTCTGTGTTCTTCAGCCAAAGA	1675
Db	708	AGGGGAAGGGGCGCTGACATCTTCATCGGCTGGGGCCCTGTCTGTGTTCTTCAGCCAAAGA	649
QY	1676	ACCGGTGGCGCTGAGTGGGGCCGTCACCTGACCCAGGAGAGAGGGCGGC	1734
Db	648	ACCGGTGGCGCTGAGTGGGGCCGTCACCTGACCCAGGAGAGAGGGCGGTCTTGGCCTC	589
QY	1735	ACGCTTCGGGGAGACTCGCTGTCTCATCGCTGCTCGCTCATTCAGGAGGCCAGGCCGCG	1794
Db	588	ACGCTTCGGGGAGACTCGCTGTCTCATCGCTGCTCATTCAGGAGGCCAGGCCGCG	529
QY	1795	GCGCTGCGCTGAAGAGGAGGCGACTACATTTGTGTCAATGATGTGAGTGAATGGCAGCCATGACGTGG	1854
Db	528	GCGCTGCGCTGAAGAGGAGGCGACTACATTTGTGTCAATGATGTGAGTGAATGGCAGCCATGACGTGG	469
QY	1855	TGGAGACACGGGAGGTGGTGAACGAGCTGAAGGCTCGGGAGAGGCGGCGCCAGGCTG	1914
Db	468	GGGAGACACGGGAGGTGGTGAACGAGCTGAAGGCTCGGGAGAGGCGGCGCCAGGCTG	409
QY	1915	CAGTGTGTGCTGTGCTGCCAGCTCTAGACTTGCCAGCTTGGGGGACCGCCGCGCCGCTC	1974
Db	408	CAGTGTGTGCTGTGCTGCCAGCTCTAGACTTGCCAGCTTGGGGGACCGCCGCGCCGCTC	349
QY	1975	CTGCTGGGCCCCAGGGGCTTCTAAGAGACGAGAGGAGCATGTTGTCAAGACCCCGGA	2034
Db	348	CTGCTGGGCCCCAGGGGCTTCTAAGAGACGAGAGGAGCATGTTGTCAAGACCCCGGA	289
QY	2035	TCCAGCTGGGCGAGTCCCGGCCCCCTCTCAACTGGAGCCGAAGGCCCHAGCAGGCGAAG	2094
Db	288	TCCAGCTGGGCGAGTCCCGGCCCCCTCTCAACTGGAGCCGAAGGCCCHAGCAGGCGAAG	229
QY	2095	ACTGGAGCTTGCCTCCAGCCCTGTGCCCCAGTGAAGCCAGCTCGGCCCTCATCCTTGAAG	2154
Db	228	ACTGGAGCTTGCCTCCAGCCCTGTGCCCCAGTGAAGCCA-STCCGCCCTCATCCTTGAAG	170
QY	2155	CACCCAGGTGGCGG	2169
Db	169	CACCCAGGTGGCGG	155

RESULT 11
BX384556
LOCUS
DEFINITION
1201 bp mRNA linear EST 08-MAY-2003
BX384556 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DK011YX10 5-PRIME, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 131 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DK011BG05Qp1.
Location/Qualifiers

```

source
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSDDK011YM10"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match      28.8%; Score 624.4; DB 13; Length 1201;
Best Local Similarity 97.3%; Pred. No. 8.2e-98;
Matches 651; Conservative 4; Mismatches 11; Indels 3; Gaps 2;

QY 468 AGCTGTCACTGTCCGCATGATCCCCCTGGGCCTGAAGGAGACCAAGAGGTGGACTGGTC 527
   |||
Db 365 AASTTACNACTGTCCCCATGATCCCCCTGGGCCTGAAGGAGACCAAGAGGTGGACTGGTC 424

QY 528 TACACCCTGAAGGAGCTGTACTCAGTGCACCTTTGGAGAGAGCGGCCCTCTCTACGAGGC 587
   |||
Db 425 TACACCCTGAAGGAGCTGTATTCAGTGCACCTTTGGAGAGAGCGGCCCTCTCTACGAGGC 484

QY 588 AGAAATCAGGAGCTGGAGGCCCTTCGGCAGGCCATCGGACCCCAAGCCGGATGAGTC 647
   |||
Db 485 AGAAATCAGGAGCTGGAGGCCCTTCGGCAGGCCATCGGACCCCAAGCCGGATGAGTC 544

QY 648 GGGCCTGGAGCTGTCTACAGCCTATTACACAGCTGTGCTTCTGATGCGCGCTTCCT 707
   |||
Db 545 GGGCCTGGAGCTGTCTACAGCCTATTACACAGCTGTGCTTCTGATGCGCGCTTCCT 604

QY 708 CACCCTGCCAGGAGCCTCGGGCTCTTCTTCCACTGTGTACGACTCGCTTACTTGGGGTCCC 767
   |||
Db 605 CACCCTGCCAGGAGCCTCGGGCTCTTCTTCCACTGTGTACGACTCGCGCTTACTTGGGGTCCC 664

QY 768 GGCCACGACGCTGCCCTTGCGCTTCGAGAAGGGCAGCGTTCTCTTCAACATCGGTGCGCT 827
   |||
Db 665 GGCCACGACGCTGCCCTTCGAGTTCGAGAGGGCAGCGTTCTCTTAAACATCGGTGCGCT 724

QY 828 CCACACGATTTGGGGCGGCCAGGACCGCTCTGCAACGAGGCTGCCCGCGCGCTAT 887
   |||
Db 725 CCACACGATTTGGGGCGGCCAGGACCGCTCTGCAACGAGGCTGCCCGCGCGCTAT 784

QY 888 GGAGGCCCTTCAGAGGGCCGCTCGGGGCTTTCAGCCTCCTCAGGAGAGAACTTCTCCCATGC 947
   |||
Db 785 GGAGGCCCTTCAGAGGGCCGCTGGGGCTTCAGCCTCCTCAGGAGAGAACTTCTCCCATGC 844

QY 948 GCGAGGCCAGACATGAGCGCTCGTCCCTCTGCGCATGAGCAGCTCATGATGGCCCA 1007
   |||
Db 845 GCGAGGCCAGACATGAGCGCTCGTCCCTCTGGCACTGAGCAGCTCATGATGGCCCA 904

QY 1008 GGCCACGGAATGTGTTTGGAGGCTCTCACCACCTGCTTCCATGGCCCCCAAGACTG 1067
   |||
Db 905 GGCCACGGAATGTGTTTGGAGGCTCTCACCACCTGCTTCCATGGCCCCCAAGACTG 963

QY 1068 CCTGGCCAGCTCGGCTGGCGCAGGAGCGGCCACAGGTGGCAGCCGAGTACAGGCTAGT 1127
   |||
Db 964 CCTTGGCCAGCTCGGCTGGCGAGAGGGGC--CCAGGTGGCAGCCGAGTACAGGCTAGT 1021

QY 1128 GCACCGGAC 1136
   |||
Db 1022 GCACSGGAC 1030

```

FEATURES

```

VERSION      BM928275.1  GI:19378654
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1. (bases 1 to 1243)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgabbs-r@mail.nih.gov
             Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLML2832 row: k column: 06
             High quality sequence start: 115
             High quality sequence stop: 630.
FEATURES     Location/Qualifiers
             1..1243
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:5770109"
                 /lab_host="DH10B"
                 /clone_lib="NIH MGC 121"
                 /notes="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
                 Site 2: EcoRV (destroyed); RNA source anonymous pool of 3
                 fetal brains, female age 20 weeks, female age 24 weeks,
                 and male age 26 weeks. Library is oligo-dT primed and
                 directionally cloned (EcoRV site is destroyed upon
                 cloning). Average insert size 1.7 kb, insert size range
                 0.7-3.5 kb. Library is normalized and enriched for
                 full-length clones and was constructed by C. Gruber
                 (Invitrogen). Research Genetics tracking code 017. Note:
                 this is a NIH_MGC Library."
ORIGIN
Query Match      28.3%; Score 613.2; DB 12; Length 1243;
Best Local Similarity 90.8%; Pred. No. 7.2e-96;
Matches 655; Conservative 0; Mismatches 63; Indels 3; Gaps 1;

QY 758 CTGGGTGCGGCGCAGACGCTGCGCTGGCCCTTCGAGAGGCGAGCGTTCTTCAACA 817
DB 27 CTGNNCCCNCCCAACACGNGGCGCTGGCCCTTCAACANGNCATCGTTCTTCAACA 86
QY 818 TCGGTGCGCTCCACACGCGATTTGGGCGCGCCAGGACCGCTCTGCACCGGGTGCCC 877
DB 87 TCNATGCCCTCCACGCAAAATTTGNGGCGCGCCANGACCACTCTCGACCGGGTGCCC 146
QY 878 GCGCGCTATGAGGCGCTTCCAGAGGCGCGCTGGGCGCTTCAGCCCTCTGAGGAGAACT 937
DB 147 GCGCGCTATGAGGCGCTTCCAGGCGCGCTGGGCGCTTCAGCCCTCTGAGGAGAACT 206
QY 938 TCTCCATGCGCGAGCCCAAGACATGAGCGCTGCTGCTCTGCGACATGAGCAGTCA 997
DB 207 TCTCCATGCGCGAGCCCAAGACATGAGCGCTGCTGCTCTGCGACATGAGCAGTCA 266
QY 998 TGAATGCCAGGCGCGAGGATGTGTTTTCAGGCGCTCTCACACCTGCTCCATGCGCC 1057
DB 267 TGAATGCCAGGCGCGAGGATGTGTTTTCAGGCGCTCTCACACCTGCTCCATGCGCC 326
QY 1058 CCAAGATGCTGCTGGCCAGCTGCGCTGCGCAGGAGGCGCCCGCAGTGGCAGCGAGT 1117
DB 327 CCAAGATGCTGCTGGCCAGCTGCGCTGCGCAGGAGGCGCCCGCAGTGGCAGCGAGT 386
QY 1118 ACAGGCTAGTGCACCGACCATGCGCCAGCAGCCGCTCCAGCTAGTGTCTCTCT 1177
DB 387 ACAGGCTAGTGCACCGACCATGCGCCAGCAGCCGCTCCAGCTAGTGTCTCTCT 446

QY 1178 GGACTGCCCTGGTGCATGTCAAGGCGGAGTACTTCGGTCTCCCTGGCCCACTACCACTAG 1237
DB 447 GGACTGCCCTGGTGCATGTCAAGGCGGAGTACTTCGGTCTCCCTGGCCCACTACCACTAG 506
QY 1238 CCATGGCCCTTGCAGCGGCTCCCGAGCGACCGAGGAGAGTCTCCACGCGAGCAGG 1297
DB 507 CCATGGCCCTTGCAGCGGCTCCCGAGCGGCTCCCGAGCGGAGAGTCTCCACGCGAGCAGG 563
QY 1298 TCTTCTGCGAGCGCCCGACCTCTCTAGCCCGGAGGCGCTGTGCTGCGCAGGAGCTGG 1357
DB 564 TCTTCTGCGAGCGCCCGACCTCTCTCTAGCCCGGAGGCGCTGTGCTGCGCAGGAGCTGG 623
QY 1358 AGGAGCGGAGGAGCTTGGCAAGGACACCTCTGAAAGCGTCCCATCTCTGGGCGAGGAGG 1417
DB 624 AGGAGCGGAGGAGCTTGGCAAGGACACCTCTGAAAGCGTCCCATCTCTGGGCGAGGAGG 683
QY 1418 CGCTGCGGCTGACGCGCTGTGCGCGTCTCTGCGGAGTGGACCTGCTGGGCGTGTGA 1477
DB 684 CGCTGCGGCGAATTTGTCGCGGCTGTCGCGGAGTGGACCTGCTGGGCGTGTGT 743
QY 1478 T 1478
DB 744 T 744

RESULT 13
LOCUS      BE251330
DEFINITION 601107871F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3344080 5',
             mRNA sequence.
ACCESSION  BE251330
VERSION    BE251330.1  GI:9121454
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1. (bases 1 to 679)
AUTHORS    NIH-MGC http://mgi.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM129 row: n column: 17
            High quality sequence stop: 617.
FEATURES   Location/Qualifiers
             1..679
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:3344080"
                 /tissue_type="retinoblastoma"
                 /lab_host="NIH MGC 16"
                 /notes="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:
                 EcoRI; cDNA made by oligo-dT priming. Directionally
                 cloned into EcoRI/XhoI sites using the following 5',
                 adaptor: GGCACGAG(G). Library constructed by Ling Hong
                 in the laboratory of Gerald M. Rubin (University of
                 California, Berkeley) using ZAP-cDNA synthesis kit
                 (Stratagene) and Superscript II RT (Life Technologies).
                 Note: this is a NIH_MGC Library."
ORIGIN
Query Match      27.9%; Score 605.6; DB 10; Length 679;
Best Local Similarity 98.1%; Pred. No. 1.1e-94;

```



```
Db      237  CTGTGCCCCAGTGAAGCCAGCTCGCGCCCTCATCTTTGAAGCAGCCAGGCTGGCG 183
|||||
RESULT 15
BG339711
LOCUS   602437292F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4555079 5',
DEFINITION mRNA sequence.
ACCESSION BG339711
VERSION   BG339711.1 GI:13146149
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1067)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ARCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI255 row: h column: 24
High quality sequence stop: 704.
Location/Qualifiers
1..1067
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4555079"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_46"
Note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
```

```

Query Match      25.6%; Score 556; DB 12; Length 1067;
Best Local Similarity 86.7%; Pred. No. 5.4e-86;
Matches 671; Conservative 0; Mismatches 95; Indels 8; Gaps 5;

QY 219 GGGCTGTGACTCCCTGACGAGATCCAGTGCGGCGCAGCTGCAGAGCCGCGAGGCCCCAGAT 278
      |||||
Db 162 GGGCTGTGACTCCCTGACGAGATCCAGTGCGGCGCAGCTGCAGAGCCGCGAGGCCCCAGAT 221
      |||||
QY 279 TCACGACGAGATTGACAAGGAGTCGAGATGCGGACGGCGCTGAGAACTCTACGAGC 338
      |||||
Db 222 TCACGACGAGATTGACAAGGAGTCGAGATGCGGACGGCGCTGAGAACTCTACGAGC 281
      |||||
QY 339 CACCAGCAACACCGGGTGAGAGAGCGTCCCTTGGAGCTGAGCTAGCTCAACTCCAA 398
      |||||
Db 282 CACCAGCAACACCGGGTGAGAGAGCGTCCCTTGGAGCTGAGCTAGCTCAACTCCAA 341
      |||||
QY 399 CTGCGAGCTGCTGAAGAGGAGCTGAGGAGCTACGCGTGGCGTGGACCCCTGGCGGCA 458
      |||||
Db 342 CTGCGAGCTGCTGAAGAGGAGCTGAGGAGCTACGCGTGGCGTGGACCCCTGGCGGCA 401
      |||||
QY 459 TGGGACGAGCTGTCACTGTCCCATGATCCCTTGGCGCTGAAGGAGCAAGGAGCT 518
      |||||
```

Search completed: July 13, 2004, 08:21:02
Job time : 3673.68 secs

Db 463 GAGACGGTCGCTCGAGCTGAGCTAGCTCAACTCCAACTGTCAGCTGCTGAGGAGGAG 522
Qy 421 CTGGAGAGCTCAGCGGTGGCGTGGACCTTGGCCGCAATGGAGCGAAGCTGTCACTGTC 480
Db 523 CTGGAGAGCTCAGCGGTGGCGTGGACCTTGGCCGCAATGGAGCGAAGCTGTCACTGTC 582
Qy 481 CCCATGATCCCCCTGGCGCTGAGGAGACCAAGAGCTGACCTGCTTACACCGCTGAAG 540
Db 583 CCCATGATCCCCCTGGCGCTGAGGAGACCAAGAGCTGACCTGCTTACACCGCTGAAG 642
Qy 541 GAGCTGATCTCAGTGCACATTTGGAGAGGACGGCGCTCTCTACGAGGAGAGAAATCAGGAG 600
Db 643 GAGCTGATCTCAGTGCACATTTGGAGAGGACGGCGCTCTCTACGAGGAGAGAAATCAGGAG 702
Qy 601 CTGGAGGCTTGGCGAGGCGCATGCGGACCCCCAGCGCGGAATGAGTGGGCTTGGAGCTG 660
Db 703 CTGGAGGCTTGGCGAGGCGCATGCGGACCCCCAGCGCGGAATGAGTGGGCTTGGAGCTG 762
Qy 661 CTCACAGCTATTACACAGCTGTGCTTCTTGATGCGCGCTTCTCTCACCCCTGCCAGG 720
Db 763 CTCACAGCTATTACACAGCTGTGCTTCTTGATGCGCGCTTCTCTCACCCCTGCCAGG 822
Qy 721 AGCTCGGGCTCTTTCTTCCACTGCTAGCTACCTGCTTACTGGGCTTCCGGCCACGACGCT 780
Db 823 AGCTCGGGCTCTTTCTTCCACTGCTAGCTACCTGCTTACTGGGCTTCCGGCCACGACGCT 882
Qy 781 GCGCTGGCTTCCAGAGGCGAGGCTTCTTCTTAAATCGGTGCTTCCACACGCGAGATT 840
Db 883 GCGCTGGCTTCCAGAGGCGAGGCTTCTTCTTAAACATCGGTGCTTCCACACGCGAGATT 942
Qy 841 GGGCGCGCCAGACGCTCTCTGACCGAGGCTGCGCGCGCTATGGAGGCTTCCAG 900
Db 943 GGGCGCGCCAGACGCTCTCTGACCGAGGCTGCGCGCGCTATGGAGGCTTCCAG 1002
Qy 901 AGGCGCGCTGGGCGCTTTCAGCTCTCTGAGGAGAACTTCTCCCATGCGCCGAGCCAGAC 960
Db 1003 AGGCGCGCTGGGCGCTTTCAGCTCTCTGAGGAGAACTTCTCCCATGCGCCGAGCCAGAC 1062
Qy 961 ATGAGCGCTGCTTCTGCGCACTGGAGAGCTCATGATGGCCGAGGCGCCAGGATGT 1020
Db 1063 ATGAGCGCTGCTTCTGCGCACTGGAGAGCTCATGATGGCCGAGGCGCCAGGATGT 1122
Qy 1021 GTGTTTTCAGGCGCTTCTCACACCTGCTTCCATGCGCCGCCCAAGACTGCTTGGCGCGCTG 1080
Db 1123 GTGTTTTCAGGCGCTTCTCACACCTGCTTCCATGCGCCGCCCAAGACTGCTTGGCGCGCTG 1182
Qy 1081 CGCTGGCGAGGAGGCGCCCGAGGTGGCGAGCGAGTACAGGTAGTGCACCGGACCATG 1140
Db 1183 CGCTGGCGAGGAGGCGCCCGAGGTGGCGAGCGAGTACAGGTAGTGCACCGGACCATG 1242
Qy 1141 GCGGAGCGCGCTGACGAGCTAGCTGCTGCTCTGGAAGTGGCGCTGCTGCAATGTCAG 1200
Db 1243 GCGGAGCGCGCTGACGAGCTAGCTGCTGCTCTGGAAGTGGCGCTGCTGCAATGTCAG 1302
Qy 1201 GCGGAGTACTTCCGCTCCCTGGCGCACTACACAGTACGATGGCGCTTCTGCGAGCGCTCC 1260
Db 1303 GCGGAGTACTTCCGCTCCCTGGCGCACTACACAGTACGATGGCGCTTCTGCGAGCGCTCC 1362
Qy 1261 CCAGCGACCGAGGAGAGCTCCCAACGAGAGAGTCTTCTGCGAGCGCCCGCCACCTCC 1320
Db 1363 CCAGCGACCGAGGAGAGCTCCCAACGAGAGAGTCTTCTGCGAGCGCCCGCCACCTCC 1422
Qy 1321 TCTAAGCGCGAGGCGCTTCTGCTGCGCGAGGAGTGGAGGAGCGAGGAGCTTGGCAAG 1380
Db 1423 TCTAAGCGCGAGGCGCTTCTGCTGCGCGAGGAGTGGAGGAGCGAGGAGCTTGGCAAG 1482
Qy 1381 GCACACTGAGAGCGCTGCTTCTTGGGAGGAGGAGGCGCTTGGCGCTGCAAGCGCTGTGC 1440
Db 1483 GCACACTGAGAGCGCTGCTTCTTGGGAGGAGGAGGCGCTTGGCGCTGCAAGCGCTGTGC 1542
Qy 1441 CGCGTCTGCGGAGGAGGAGCTTCTGCGGCTGTGATCTCCAGAGCTGCGAGGCTCA 1500
Db 1543 CGCGTCTGCGGAGGAGGAGCTTCTGCGGCTGTGATCTCCAGAGCGCTGCGAGGCTCA 1602

Qy 1501 CTGGCCAACTATGCGGAGCTCGACCGTGGAGATGACTTCTGTGAGGCTGCCGAGGCCCG 1560
Db 1603 CTGGCCAACTATGCGGAGCTCGACCGTGGAGATGACTTCTGTGAGGCTGCCGAGGCCCG 1662
Qy 1561 GACATCCAGCTTAAGACCCACCAAGAACCCAGAGGCGCAGGATGCCACGCTGTCCAGGGG 1620
Db 1663 GACATCCAGCTTAAGACCCACCAAGAACCCAGAGGCGCAGGATGCCACGCTGTCCAGGGG 1722
Qy 1621 AAGGGCGCTGACATCTTCATCGGCTGGGGCCCTCTGTGTCTTCTCAGCCAGAGACCGG 1680
Db 1723 AAGGGCGCTGACATCTTCATCGGCTGGGGCCCTCTGTGTCTTCTCAGCCAGAGACCGG 1782
Qy 1681 TGGCGGCTGTGTGGGCGCTTCCACTGACCCGAGGAGAGGGCGGCTTGTGGCTTCAAGCTT 1740
Db 1783 TGGCGGCTGTGTGGGCGCTTCCACTGACCCGAGGAGAGGGCGGCTTGTGGCTTCAAGCTT 1842
Qy 1741 CGGGGAGACTCGCTGTCTCTCATCGCTGCGCTCATTCAGAGGAGCCAGGCGCGGCGGCT 1800
Db 1843 CGGGGAGACTCGCTGTCTCTCATCGCTGCGCTCATTCAGAGGAGCCAGGCGCGGCGGCT 1902
Qy 1801 GGCCTGAAGAGGCGGCTTACATTTGTCTAGTGAATGGGCGAGCCATGSCAGGTGCTGAGG 1860
Db 1903 GGCCTGAAGAGGCGGCTTACATTTGTCTAGTGAATGGGCGAGCCATGSCAGGTGCTGAGG 1962
Qy 1861 CACGCGAGGTGTGTGAGGAGCTGAAGGCTTGGCGGAGAGGCGGCGCCAGCTTGCAGGTG 1920
Db 1963 CACGCGAGGTGTGTGAGGAGCTGAAGGCTTGGCGGAGAGGCGGCGCCAGCTTGCAGGTG 2022
Qy 1921 GTGTGCTGTGCTGCCAGCTTACAGTCCAGCTTGGGGAGACCGCGCGGCGCTTGTGCTG 1980
Db 2023 GTGTGCTGTGCTGCCAGCTTACAGTCCAGCTTGGGGAGACCGCGCGGCGCTTGTGCTG 2082
Qy 1981 GGCCTGAGGCGGCTTCTTACGAGCCAGAGGAGCATGTTGCAAGACCCCGGCTTCCAGG 2040
Db 2083 GGCCTGAGGCGGCTTCTTACGAGCCAGAGGAGCATGTTGCAAGACCCCGGCTTCCAGG 2142
Qy 2041 TGGCGCACTCCCGGCGCTTCTCAACTGGAGCGGAAAGCCAGCAGGCGGCAAGACTGGA 2100
Db 2143 TGGCGCACTCCCGGCGCTTCTCAACTGGAGCGGAAAGCCAGCAGGCGGCAAGACTGGA 2202
Qy 2101 GGTGCGCCCGCAGCGCTTGTGCGCCAGTGAAGCGAGCTTCCGCGCTTCTGTAAGCACCCA 2160
Db 2203 GGTGCGCCCGCAGCGCTTGTGCGCCAGTGAAGCGAGCTTCCGCGCTTCTGTAAGCACCCA 2262
Qy 2161 GGTGCGCG 2169
Db 2263 GGTGCGCG 2271

RESULT 2

US-10-274-878-1

; Sequence 1, Application US/10274878

; Patent No. 6670163

; GENERAL INFORMATION:

; APPLICANT: RUSCH, Douglas et al

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CL001099-CIP-DIV

; CURRENT APPLICATION NUMBER: US/10/274,878

; CURRENT FILING DATE: 2002-10-22

; PRIOR APPLICATION NUMBER: 09/849,334

; PRIOR FILING DATE: 2001-05-07

; PRIOR APPLICATION NUMBER: 09/773,371

; PRIOR FILING DATE: 2001-02-01

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: Fast-Seq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 2469

; TYPE: DNA

; ORGANISM: Human

US-10-274-878-1

Query Match 100.0%; Score 2169; DB 4; Length 2469; Best Local Similarity 100.0%; Pred. No. 0; Matches 2169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;												
QY	1	ATGATCTCTGGAGGAGAGCGCGACGCGCGGGCGCGGAGGAGAGCCCGCGCTGCAG	60						QY	1021	GTGTTTTCAGGGGCTCTCACACCTGCTCCTCATGGCCCCCAAGACTGCTCTGGCCCCAGCTG	1080
DB	103	ATGATCTCTGGAGGAGAGCGCGACGCGCGGGCGCGGAGGAGAGCCCGCGCTGCAG	162						DB	1123	GTGTTTTCAGGGGCTCTCACACCTGCTCCTCATGGCCCCCAAGACTGCTCTGGCCCCAGCTG	1182
QY	61	ATATCTAGGAGAAAAACCCAGGAAAAACAGTGTGAGCTCTTTACGGGGAAGACCGGAAGGC	120						QY	1081	CGCTGGCGCAGGAGGCGGCCAGGTGGCAGCCGAGTACAGGCTAGTGACCCGGAACCATG	1140
DB	163	ATATCTAGGAGAAAAACCCAGGAAAAACAGTGTGAGCTCTTTACGGGGAAGACGGGAAGGC	222						DB	1183	CGCTGGCGCAGGAGGCGGCCAGGTGGCAGCCGAGTACAGGCTAGTGACCCGGAACCATG	1242
QY	121	CTGAGAGACGTGTGTGCGTGGAGAGGGTGTGCGGTCCACAGAGGGGAAGACCCAGTGCCT	180						QY	1141	GCCAGACCAACCGTCCACGACTACGTGCTGTCTCTCTGGAGCTGCCCCCTGGTGCAATGTCAAG	1200
DB	223	CTGAGAGACGTGTGTGCGTGGAGAGGGTGTGCGGTCCACAGAGGGGAAGACCCAGTGCCT	282						DB	1243	GCCAGACCAACCGTCCACGACTACGTGCTGTCTCTCTGGAGCTGCCCCCTGGTGCAATGTCAAG	1302
QY	181	GTGCACGTTGGCCCCCATGAATCCGACGCTTCATGACGTGGGCTGTGACTCCCTGACGCGAG	240						QY	1201	GCCGAGTACTTTCGCTCCTTGGCCCCACTACCACTAGCCATGGCCCCCTCTGCGACGGCTCC	1260
DB	283	GTGCACGTTGGCCCCCATGAATCCGACGCTTCATGACGTGGGCTGTGACTCCCTGACGCGAG	342						DB	1303	GCCGAGTACTTTCGCTCCTTGGCCCCACTACCACTAGCCATGGCCCCCTCTGCGACGGCTCC	1362
QY	241	ATCCAGTGGCCAGCTGCAGAGCCGACGGCCCCAGATTCCACGACAGATTGACAAAGGAG	300						QY	1261	CCAGCGACCGAGGAGAGCTCCCGACGACGAGAGAGGTCTTCTGAGAGCCCCCACCCTCC	1320
DB	343	ATCCAGTGGCCAGCTGCAGAGCCGACGGCCCCAGATTCCACGACAGATTGACAAAGGAG	402						DB	1363	CCAGCGACCGAGGAGAGCTCCCGACGACGAGAGAGGTCTTCTGAGAGCCCCCACCCTCC	1422
QY	301	CTGCAGATCGGACGGGCGCTGAAACCTCTACAGAGCCACAGCAACACCGGGTGAGA	360						QY	1321	TCTAAGCCCCCAGGCGCTGTGCTGCGCAGGAGCTGGAGAGCGCAGGAGCTTGGCAAG	1380
DB	403	CTGCAGATCGGACGGGCGCTGAAACCTCTACAGAGCCACAGCAACACCGGGTGAGA	462						DB	1423	TCTAAGCCCCCAGGCGCTGTGCTGCGCAGGAGCTGGAGAGCGCAGGAGCTTGGCAAG	1482
QY	361	GAGACGCTCGCCTGGAGCTGAGCTACGTCACCTCCAACTCGACGCTGCTGAAGAGGAG	420						QY	1381	GCACACTGAAGCGTGCTATCTTGGGGCAGGAGAGCGCTGCGGCTGCACGCCCTGTGTC	1440
DB	463	GAGACGCTCGCCTGGAGCTGAGCTACGTCACCTCCAACTCGACGCTGCTGAAGAGGAG	522						DB	1483	GCACACTGAAGCGTGCTATCTTGGGGCAGGAGAGCGCTGCGGCTGCACGCCCTGTGTC	1542
QY	421	CTGAGAGAGCTCAGCGGTGGGTGGAGCCCTGCGCGCGCATGGAGCGGAAGCTGTCACTGTC	480						QY	1441	CGGTCCTGCGCGAGGTGGACCTGTTCCGGGCTGTGATCTCCAGAGCTGTCAGCGCTCA	1500
DB	523	CTGAGAGAGCTCAGCGGTGGCGTGGAGCCCTGGCGCGCATGGAGCGGAAGCTGTCACTGTC	582						DB	1543	CGGTCCTGCGCGAGGTGGACCTGCTTCCGGGCTGTGATCTCCAGAGCTGTCAGCGCTCA	1602
QY	481	CCCATGATCCCTCGGCGCTGAGGAGACCAAGAGAGCTGGACTGGTCTACACCGCTCAAG	540						QY	1501	CTGGCCAAAGTATCGGAGCTCGACCGGTGAGGATGACTTCTGTGAGGCTGCCAGAGGCCCG	1560
DB	583	CCCATGATCCCTCGGCGCTGAGGAGACCAAGAGAGCTGGACTGGTCTACACCGCTCAAG	642						DB	1603	CTGGCCAAAGTATCGGAGCTCGACCGGTGAGGATGACTTCTGTGAGGCTGCCAGAGGCCCG	1662
QY	541	GAGCTGATCTCAGTGCACTTTGGAGAGGACGGCGCTCTCTACAGGAGCAAAATCAGGGAG	600						QY	1561	GACATCCAGCCTTAAGACCCACAGAGACCCAGAGGCCAGGATGCCAGCTGTCCGAGGGG	1620
DB	643	GAGCTGATCTCAGTGCACTTTGGAGAGGACGGCGCTCTCTACGAGGACAGAAATCAGGGAG	702						DB	1663	GACATCCAGCCTTAAGACCCACAGAAAGCCAGAGGCCAGGATGCCAGCGCTGTCCGAGGGG	1722
QY	601	CTGAGGCGCTCGGCGAGCCATGCGGACCCCGCAGCCGGAATGAGTTCGGGCTTGAGCTG	660						QY	1621	AAGGGGCTTGACATCTTTCATCGGCTGGGGCCCCCTGTCTGTGTTCTCAGCCAAAGAACCG	1680
DB	703	CTGAGGCGCTTCGCGCAGGCGCATGCGGACCCCGCAGCCGGAATGAGTTCGGGCTTGAGCTG	762						DB	1723	AAGGGGCTTGACATCTTTCATCGGCTTGGGGCCCCCTGTCTGTGTTCTCAGCCAAAGAACCG	1782
QY	661	CTCAGAGCTTATCAACAGAGCTGTGCTTCTGGAATGCGCGCTTCTCACCCCTGCCAGG	720						QY	1681	TGGCGGCTGGTGGGGCCCGTCCACCTGACCCGAGGAGAGGGCGGCTTTGGCCTCAGGCTT	1740
DB	763	CTCAGAGCTTATCAACAGAGCTGTGCTTCTTGAATGCGCGCTTCTCACCCCTGCCAGG	822						DB	1783	TGGCGGCTGGTGGGGCCCGTCCACCTGACCCGAGGAGAGGGCGGCTTTGGCCTCAGGCTT	1842
QY	721	AGCCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTACTGGGGTCCCGGCCAGCAGCGT	780						QY	1741	CGGGGAGACTCGGCTGTCTCATCGCTGCGCTCAATTCCAGGGAGCCAGGCCCGGGCGGCT	1800
DB	823	AGCCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTACTGGGGTCCCGGCCAGCAGCGT	882						DB	1843	CGGGGAGACTCGGCTGTCTCATCGCTGCGCTCAATTCCAGGGAGCCAGGCCCGGGCGGCT	1902
QY	781	GCCCTGGCTTCGAGAAAGGACGCTTCTTCAACATCGGTGCGCTCCACACGAGATT	840						QY	1801	GCGCTGAAGGAGGGCGACTACATTTGTGTGCTAGTGAATGGGAGCCATGAGGTGTGGAGA	1860
DB	883	GCCCTGGCTTCGAGAAAGGACGCTTCTTCAACATCGGTGCGCTCCACACGAGATT	942						DB	1903	GCGCTGAAGGAGGGCGACTACATTTGTGTGCTAGTGAATGGGAGCCATGAGGTGTGGAGA	1962
QY	841	GGGGCGCGCAGGACCGCTCTGCAACGAGGTGCGCGCTGATGAGGCGCTTCCAG	900						QY	1861	CACGCGAGGTGTGACGAGCTGAAAGCTGCGGAGAGGGCGGGCGCAGCTGTCAGGTG	1920
DB	943	GGGGCGCGCAGGACCGCTCTGCAACGAGGTGCGCGCTGATGAGGCGCTTCCAG	1002						DB	1963	CACGCGAGGTGTGTCGAGCTGAAAGCTGCGGAGAGGGCGGGCGCAGCTGTCAGGTG	2022
QY	901	AGGGCGCTGGGGCTTCAGCCCTCTGAGGGAGAACTTCTCCCATGCGCGGAGCCAGAC	960						QY	1921	GTGTGCTCTGCCCCAGCTCTAGACTTGCCAGCTTGGGGGACCCCGGCCCTGCTCTGCTG	1980
DB	1003	AGGGCGCTGGGGCTTCAGCCCTCTGAGGGAGAACTTCTCCCATGCGCGGAGCCAGAC	1062						DB	2023	GTGTGCTCTGCCCCAGCTCTTAGACTGCCACGCTTGGGGGACCCCGCGGCCCTGCTCTGCTG	2082
QY	961	ATGAGCGCTGCGTCCCTCTGCGCACTGGAGCAGCTCATGATGGCCCGAGGCCAGGAATGT	1020						QY	1981	GGCCCCCAGGGGCTTCTTAAGAGCCAGAGGGAGCATGTTGCAAGACCCCGGCACTCCAG	2040
DB	1063	ATGAGCGCTGCGTCCCTCTGCGCACTGGAGCAGCTCATGATGGCCCGAGGCCAGGAATGT	1122						DB	2083	GGCCCCCAGGGGCTTCTAAGGAGCCAGAGGGAGCATGTTGCAAGACCCCGGCACTCCAG	2142
									QY	2041	TGGGCGAGTCCCGGCCCTCTCAACTGGAGCCGAAAGGCCAGAGCGGCAAGACTGGA	2100
									DB	2143	TGGGCGAGTCCCGGCCCTCTCAACTGGAGCCGAAAGGCCAGAGCGGCAAGACTGGA	2202
									QY	2101	GGCTGCCCGCCAGCCCTGTGCCCCAGTGAAGCAGCTCCGCCCTCATCTTGAAGCACCCA	2160

Db 2203 GGCTGCCCCAGCCCTGTGCCCCAGTGAAGCCAGCTCCGCCCTCACTCTTGAAGACCCCA 2262
QY 2161 GGCTGGCCG 2169
Db 2263 GGCTGGCCG 2271

RESULT 3
US-09-833-381-1293
; Sequence 1293, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: ROBISON, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1293
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1155)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1293

Query Match 36.3%; Score 787; DB 4; Length 1155;
Best Local Similarity 99.4%; Pred. No. 4e-156;
Matches 790; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1375 GGCAAGGCACACCTGAAGCGTGCCATCTCTGGGGCAGAGGAGGCGCTGCGCTGACGCC 1434
Db 133 GGCAGGCACCTCTGAAGCGTGCCATCTCTGGGGCAGAGGAGGCGCTGCGCTGACGCC 192
QY 1435 CTGTGCGCGCTCTCTGCGGAGTGAGCTGTCTTGGGCTGTGATCTCTCCAGAGCTGCG 1494
Db 193 CTGTGCGCGCTCTCTGCGGAGTGAGCTGTCTTGGGCTGTGATCTCTCCAGAGCTGCG 252
QY 1495 CGCTCACTGGCCAAATGCGGAGCTGACCGTGAAGATGATCTCTGAGGCTGCGGAG 1554
Db 253 CCCCACCGCGCAAGTATGCGGAGCTGACCGTGAAGATGATCTCTGAGGCTGCGGAG 312
QY 1555 GCCCGGACATCCAGCCTTAAGACCCACAGAGCCAGAGGCGCAGATGCCACGCTGTCC 1614
Db 313 GCCCGGACATCCAGCCTTAAGACCCACAGAGCCAGAGGCGCAGATGCCACGCTGTCC 372
QY 1615 CAGGGGAAGGGGCTGACATCTTCCATGCGTGGGGCCCTGTCTGTGTTCTAGCCAAAG 1674
Db 373 CAGGGGAAGGGGCTGACATCTTCCATGCGTGGGGCCCTGTCTGTGTTCTAGCCAAAG 432
QY 1675 AACCGTGGCGGCTGTGTGGGCGCGCTGACCTGACCCGAGAGAGGGGGTGTGGCTC 1734
Db 433 AACCGTGGCGGCTGTGTGGGCGCGCTGACCTGACCCGAGAGAGGGGGTGTGGCTC 492
QY 1735 ACGCTTCGGGAGACTGCGCTGTCTCATCGCTGCGCTCATTCAGGGAGGAGCGCGCG 1794
Db 493 ACGCTTCGGGAGACTGCGCTGTCTCATCGCTGCGCTCATTCAGGGAGGAGCGCGCG 552
QY 1795 GCGGCTGCGCTGAAGAGGGGCGACTACATTTGTGTAGTGAATGGGAGGCCATGAGGTGG 1854
Db 553 GCGGCTGCGCTGAAGAGGGGCGACTACATTTGTGTAGTGAATGGGAGGCCATGAGGTGG 612
QY 1855 TGGAGACACCGGGAGGTGTGACGAGCTGAGAGCTGGGAGAGGGGGCGCCAGCTG 1914
Db 613 TGGAGACACCGGGAGGTGTGACGAGCTGAGAGCTGGGAGAGGGGGCGCCAGCTG 672
QY 1915 CAGGTGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTTGGGGGACCGCGCGCCGCTC 1974

Db 673 CAGGTGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTTGGGGAGACCGCGCCGCTC 732
QY 1975 CTGCTGGGCCCCCAGGGGGCTTTTAAGAGCCAGAGGAGCATGTTGCAAGACCCCGCA 2034
Db 733 CTGCTGGGCCCCCAGGGGGCTTTTAAGAGCCAGAGGAGCATGTTGCAAGACCCCGCA 792
QY 2035 TCCACGTGGGCGCAGTCCCGGGCCCTCTCAACTGGAGCCGAAAGGCCAGAGGCGCAAG 2094
Db 793 TCCACGTGGGCGCAGTCCCGGGCCCTCTCTCACTGGAGCCGAAAGGCCAGAGGCGCAAG 852
QY 2095 ACTGAGAGGCTGCCCCCAGGCCCTGTGCCCCAGTGAAGCCAGCTCGGCCCTCATCTTGAAG 2154
Db 853 ACTGAGAGGCTGCCCCCAGGCCCTGTGCCCCAGTGAAGCCAGCTCGGCCCTCATCTTGAAG 912
QY 2155 CACCAGAGGTGGCGG 2169
Db 913 CACCAGAGGTGGCGG 927

RESULT 4
US-09-849-334-3
; Sequence 3, Application US/09849334
; Patent No. 6500655
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001099-CIP
; CURRENT APPLICATION NUMBER: US/09/849,334
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 19025
; TYPE: DNA
; ORGANISM: Human
US-09-849-334-3

Query Match 10.0%; Score 217; DB 4; Length 19025;
Best Local Similarity 76.6%; Pred. No. 2e-36;
Matches 321; Conservative 0; Mismatches 10; Indels 88; Gaps 1;
QY 1638 CCATCGGCTGCGGCGCCCTGTCTGTGTTCTCAGCCAGAAACCGGTGGCGGTGTTGGGGCC 1697
Db 14938 CCGCTGCCAGGGGCGCCCTGTCTGTGTTCTCAGCCAGAAACCGGTGGCGGTGTTGGGGCC 14997
QY 1698 CGTCCACTGAACCGGAGAGGCGGCTTTGGCTTTCAGCTTCGGGGAGACTGCGCTGT 1757
Db 14998 CGTCCACTGAACCGGAGAGGCGGCTTTGGCTTTCAGCTTCGGGGAGACTGCGCTGT 15057
QY 1758 CTTCACTGCTGCGCTCATTCAGGGAGCCAGGCGC----- 1793
Db 15058 CTTCACTGCTGCGCTCATTCAGGGAGCCAGGCGCGGTGAAGGCGCGCGCGCGCTG 15117
QY 1794 ----- 1793
Db 15118 AGGCTGAGTCTTGGTCCAGCAGAGGTGTCTGTCCCCACCTCACCGTCCAACTCTCCC 15177
QY 1794 ----GGCGGCTGGCTTGAAGAGGGCGACTACATTTGTGTGATGATGGGAGCGCATGCA 1849
Db 15178 CACAGGCGGCTGGCTTGAAGAGGGCGACTACATTTGTGTGATGATGGGAGCGCATGCA 15237
QY 1850 GGTGCTGGAGACACCGGAGGTGTGACGAGGCTGAAGGCTTGGGGAGAGCGCGGCGCA 1909
Db 15238 GGTGCTGGAGACACCGGAGGTGTGACGAGCTGAAGGCTTGGGGAGAGCGCGGCGCA 15297
QY 1910 GCCTGCAAGGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGAGACCGCGCG 1968
Db 15298 GCCTGCAAGGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGAGACCGCGCG 15356

```
RESULT 5
US-10-274-878-3
; Sequence 3, Application US/10274878
; Patent No. 6670163
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001099-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,878
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 19025
; TYPE: DNA
; ORGANISM: Human
US-10-274-878-3

Query Match
Best Local Similarity 10.0%; Score 217; DB 4; Length 19025;
Matches 321; Conservative 0; Mismatches 10; Indels 88; Gaps 1;

QY 1638 CCATCGCTGGGGCCCTGTCTGTGTTCTCAGCCCAAGAACCGGTGGCGCTGTGGGGCC 1697
Db 14938 CCCCTGCCAGGGGCCCTGTCTGTGTTCTCAGCCCAAGAACCGGTGGCGCTGTGGGGCC 14997

QY 1698 GTCACACCTGACCCGAGAGAGAGGGCGCTTGGCTCAGCTTCGGGGAGACTCGCTGT 1757
Db 14998 GGTCCACCTGACCCGAGAGAGAGGGCGCTTGGCTCAGCTTCGGGGAGACTCGCTGT 15057

QY 1758 CTTATCGCTGCGCTCATTTCCAGGAGGAGCCAGCCGCGC----- 1793
Db 15058 CTTATCGCTGCGCTCATTTCCAGGAGGAGCCAGCCGCGCGTAAGGGCCCGCGCCCTGT 15117

QY 1794 ----- 1793
Db 15118 AGGCTGAGTCTTGTGTCAGCAGCGGTGCTCTGTCCTCCCACTCAGCGTCCAGTCTCC 15177

QY 1794 -----GGCGCTGCTTGAAGAGGGCGACTACATTGTGTCACTGAATGGCAGCATGCA 1849
Db 15178 CACAGGCGCTGGCTGAGGAGGGCGACTACATTGTGTCACTGAATGGCAGCATGCA 15237

QY 1850 GGTGTGGAGACACCGCGAGGTGGTGACGAGCTGAAGCTCAGGAGGGCGGCGGCGCA 1909
Db 15238 GGTGTGGAGACACCGAGAGGTGTGACGAGCTGAAGCTCAGGAGGGCGGCGGCGCA 15297

QY 1910 GCCTCAGGTGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACCCGCG 1968
Db 15298 GCCTCAGGTGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGTGAGCCCTGTG 15356

RESULT 6
US-09-849-334-1/c
; Sequence 1, Application US/09849334
; Patent No. 6500655
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001099-CIP
; CURRENT APPLICATION NUMBER: US/09/849,334
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Human
US-10-274-878-1
```

```
US-10-274-878-1/c
; Sequence 1, Application US/10274878
; Patent No. 6670163
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001099-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,878
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Human
US-10-274-878-1

Query Match
Best Local Similarity 3.1%; Score 68; DB 4; Length 2469;
Matches 158; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 1344 GCCGAGGAGCTGGAGGAGCGCAGCTTGGCAAGGCACACCTGAAGCGTGCCATCCT 1403
Db 1753 GCCCAGCGATGGAAGATGTGAGGCCCTTCCCTGGGACAGCGCTGGCATCTGGCCT 1694

QY 1404 GGGGAGGAGGAGCGCTGCGCTGCACGCCCTGTGTCGCCGCTCTCGGAGGTGACCT 1463
Db 1693 CTGGCTTCTGTGGGTCTTAGGCTGGATGTCCGGGCGCTCGGAGCTCACAGAAGTCAT 1634

QY 1464 GCTTCGGGCTGTGATCTCCAGACGCTGCAGCGCTCACTGGCCAGATATGGGAGTCTGA 1523
Db 1633 CTTCAGGTCGAGCTCCGATACTTTGGCCAGTGGAGCTGTGGGAGATCACAG 1574

QY 1524 CCGTGAGGATGACTTCTGTGAGGCTGCCGAGGCCCGGACATCCAGCTTAAGACCCACCA 1583
Db 1573 CCGAAGCAGGTCCACCTCGCGAGACGCGGACAGGGCTGGAGCCCGGCTCTCT 1514

QY 1584 GAAGCAGAGGCGCAGATGCCAGCTGTCCAGGCTGTCCAGGGGAGGGGCTTGACATCTTCCATCG 1643
Db 1513 CCGGCCAGGATGCAAGCTTTCAGGTGTGCTTGCAGGCTGCTTGCAGCTCTCTCCAGCT 1454

QY 1644 GCTGGGGC 1651
Db 1453 CCTGGGC 1446

RESULT 7
US-10-274-878-1/c
; Sequence 1, Application US/10274878
; Patent No. 6670163
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001099-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,878
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Human
US-10-274-878-1

Query Match
Best Local Similarity 3.1%; Score 68; DB 4; Length 2469;
Matches 158; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 1344 GCCGAGGAGCTGGAGGAGCGCAGCTTGGCAAGGCACACCTGAAGCGTGCCATCCT 1403
Db 1753 GCCCAGCGATGGAAGATGTGAGGCCCTTCCCTGGGACAGCGCTGGCATCTGGCCT 1694

QY 1404 GGGGAGGAGGAGCGCTGCGCTGCACGCCCTGTGTCGCCGCTCTCGGAGGTGACCT 1463
Db 1693 CTGGCTTCTGTGGGTCTTAGGCTGGATGTCCGGGCGCTCGGAGCTCACAGAAGTCAT 1634

QY 1464 GCTTCGGGCTGTGATCTCCAGACGCTGCAGCGCTCACTGGCCAGATATGGGAGTCTGA 1523
Db 1633 CCTCAGGTCGAGCTCCGATACTTTGGCCAGTGGAGCTGTGGGAGATCACAG 1574
```


[illegible]

Qy	1391	AGCGTGCATCTCGGGCAGGAGGCGCTGCGGTGCACGCCTGTGCGCGCTCTGC	1450
Db	22211	GGCGCGCGCGCTCTCGCGTGGCTGAGCCAGCGCCCACTTGACCCGCGAGGCTC	22270
Qy	1451	GCAGGTGGACCTGCTTCGGGCTGTGATCTCCAGACGCTGACGCGCTCACTGGCCAAAT	1510
Db	22271	TCGCCGAGCACTTGCCTGCGCAGGCTTGCGCGAGACCGCGCGCTCGCGCGTGTCTCGC	22330
Qy	1511	ATCGGAGCTCGACCGTGAGGATGATCTTGTGAGCTGCCAGGCGCCCGGACATC	1566
Db	22331	TCCTGCGCTTCGACGAAAGTCCCTTCGCGGACCATGCGCGCTGCGCGGAGCTC	22386

```

RESULT 12
US-08-457-646A-6
? Sequence 6, Application US/08457646A
? Patent No. 5679560
? GENERAL INFORMATION:
? APPLICANT: Schupp, Thomas
? APPLICANT: Ligon, James M.
? APPLICANT: Beck, James Joseph
? APPLICANT: Hill, Dwight Steven
? APPLICANT: Ryals, John Andrew
? APPLICANT: Gaffney, Thomas Deane
? APPLICANT: Lam, Stephen Iing
? APPLICANT: Hammer, Phillip E.
? APPLICANT: Uknes, Scott Joseph
? TITLE OF INVENTION: Genes for the synthesis of
? TITLE OF INVENTION: antipathogenic substances
? NUMBER OF SEQUENCES: 22
? CORRESPONDENCE ADDRESS:
?

```

Qy	971	CGTCCCTCTGCGCATGAGACGCTCATGATGCCAGGCCAGGAATCTGTGTTGAGG	1030
Db	21797	CGACGATCCTGCCGCCCGGAAAGCGGTGCGCTGCCACCTTACCAGCGCGAGC	21856
Qy	1031	GCCTCTCACACCTGCTCTCATGGCCCCCAAGACTGCTGGCCAGCTGCGCTGCGCGC	1090
Db	21857	GCTTCTGGCTCGACGCTCTCACGGCACCGCGCGCGGCGTCAACACACTTGTCTCGCTCG	21916
Qy	1091	AGGAGGCGCGCCAGGTGGCAGCGCGAGCTACAGGCTAGTGACACGGACCATGGCCCCAGCCAC	1150
Db	21917	AGGGCGGTTCTGCGAGGCCATCGAGCGGGNAATCGACGGCTCAGCGGCCAGCTCC	21976
Qy	1151	CCGTTCAGACATGATGTCCTGTCTCTGAGCTGCCTGTGTGATGTCAAGGCCAGTACT	1210
Db	21977	ACGTGAGCGGACGACGAGCGCGCGCCCTTGCCCTGCTCTTCCACCCCTCGCGAGCT	22036
Qy	1211	TCGCTCTCCCTGGCCCCACTACCACTAGACCATGGCCCTCTGCGACGGCTCCCAGCGACGG	1270
Db	22037	TTCCGCACGAGCGGCAAGAGCAGGACCGGTGACGCGCTTGGCGCTACCGCATCACTGGA	22096
Qy	1271	AGGAGAGCTTCCCAACGACAGCAGAGGTCTTCTGTGAGCCCCCACCCTCTCTAAGCCCC	1330
Db	22097	AGCCTCTGACCAACGCCACACCGCCCGCCACTGG-----CCGGCACTTGGCTCTCTCG	22150
Qy	1331	GAGGCCCTGTGCTGCCGACGAGCTGAGAGGAGCGAGGAGGCTTGGCAAGGCACACCTGA	1390
Db	22151	TCGTGCGGCGCGCTCTGGACGACGACGCGCTCCCCCTCGCGGTCAACGAGGCGCTCGGCC	22210
Qy	1391	AGCGTGCCATCTCTGGGGCAGGAGGCGCTGGGCTGCAGCCCTGTGCCGCGCTCTGCG	1450
Db	22211	GGCGGCGGCGCGCTCTCGCGCTGCGCTTGGCGAGGCCACCTGACCCGCGAGGCTC	22270
Qy	1451	GCAGGTGGACCTGTCTTCGGGCTGTGATCTCCGACAGCTGCAGCGCTCACTTGGCCAAGT	1510
Db	22271	TCGCCGAGCACTTGCACAGGCTTGCGCGAGACCGCGCGCTCGCGGCTGTCTCTCGC	22330
Qy	1511	ATCGGAGCTCGACCTGTAGGATGATCTTGTGAGGCTGCCAGGCCCCGACATC	1566
Db	22331	TCTCGCCCTTCGACGAAAGTCCCTTCGCGGACCATGTGCGCGCTGCCGCGGATC	22386

```

RESULT 13
US-08-458-076A-6
; Sequence 6, Application US/08458076A
; Patent No. 5698425
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/458,076A
; APPLICATION NUMBER: US/08/458,076A
; FILING DATE: 01-JUN-1995

```


Db	22037	TTCCGCACGAGCGCGCAGAGAGGCGACGGTCGACGCTGGCGGCTACCGCATCAGCTGGA	22096
Qy	1271	AGGAGAGCTCCCCACGCACGACGACGAGGTCTTCCTGCAGCCCCCACTCTCTCTAAGCCCC	1330
Db	22097	AGCCTCTGACACCGGCACACGCCGCGACCTGG-----CGGCACTGGCTCCTCG	22150
Qy	1331	GAGGCCCTGTGTCGCCGACGAGAGCTGGAGGAGCGAGGCAGCTTGGCAGGACACACTGA	1390
Db	22151	TGCTGCGGCCGCTCTGACGACGCGCTCCCTCCGCGCTACCGAGGCGCTCGGCC	22210
Qy	1391	AGGTCGCATCTTGGGGCAGGAGGCGCTCGGCTGCAGCCCTGTGCCGCGCTCTCTGC	1450
Db	22211	GCGCGCGCGCGCGCTCTCGCGCTGCGCTGAGCGAGCCACCTGAGACCGAGGCTC	22270
Qy	1451	GCGAGGTGACCTGCTTCGGGCTGTATCTCCAGACGCTCGAGGCTCACTGCCCAAGT	1510
Db	22271	TGCGGAGCACTTGCGCAGGCTTGGCGCGAGACGCGCGCTCGGCGGTGCTCTGC	22330
Qy	1511	ATGCGGAGCTCGACCGTGAGGATGCTTGTGAGGCTGCCGAGCCCCGACATC	1566
Db	22331	TCTCGCCCTCGAGAAAGTCCCTCGCCGACCATGCGCGCGTGCCTCGCGGAGCTC	22386

RESULT 15

```

US-08-457-335A-6
/ Sequence 6, Application: US/08457335A
/ Patent No. 5723759
/ GENERAL INFORMATION:
/ APPLICANT: Schupp, Thomas M.
/ APPLICANT: Ligon, James M.
/ APPLICANT: Beck, James Joseph
/ APPLICANT: Hill, Dwight Steven
/ APPLICANT: Ryals, John Andrew
/ APPLICANT: Gaffney, Thomas Deane
/ APPLICANT: Lam, Stephen Ting
/ APPLICANT: Hammer, Phillip E.
/ APPLICANT: Uknes, Scott Joseph
/ TITLE OF INVENTION: Genes for the synthesis of
/ TITLE OF INVENTION: antipathogenic substances
/ NUMBER OF SEQUENCES: 22
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Ciba-Geigy Corporation
/ STREET: 7 Skyline Drive
/ CITY: Hawthorne
/ STATE: NY
/ COUNTRY: USA
/ ZIP: 10512
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/457,335A
/ FILING DATE: 01-JUN-1995
/ CLASSIFICATION: 800
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/457,205
/ FILING DATE: 01-JUN-1995
/ APPLICATION NUMBER: 08/258,261
/ FILING DATE: 08-Jun-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Elmer, James Scott
/ REGISTRATION NUMBER: 36,129
/ REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 919-541-8614
/ TELEFAX: 919-541-8689
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2958 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2004, 19:58:24 ; Search time 559.787 Seconds
(without alignments)
16460.437 Million cell updates/sec

Title: US-10-697-266-1_COPY_103_2271

Perfect score: 2169

Sequence: 1 atgacctggaggagagcc.....tgaagcaccacgggtggccg 2169

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 29Jan04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002s:*

7: Geneseqn2003as:*

8: Geneseqn2003bs:*

9: Geneseqn2003cs:*

10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2169	100.0	2469	6	AAL48837
2	1951	89.9	2837	7	ABV72511
3	798.8	36.8	3213	7	ACC46698
4	781.4	36.0	1013	5	AAS00831
5	779.8	36.0	1013	3	AAC77780
6	496.4	22.9	2100	5	AAL15828
7	486.4	22.9	3526	6	AAS15827
8	494.8	22.8	2109	6	ABX97183
9	493.2	22.7	2061	6	ABX71856
10	493.2	22.7	3484	6	ABX71855
11	487.8	22.5	3019	4	AAF58361
12	463.6	21.4	2310	6	ABX97182
13	436	19.3	764	2	AAZ15546
14	418.2	19.3	2757	4	AAAC91349
15	379.2	17.5	1671	6	ABK92255
16	355	16.4	3335	5	AAS72556
17	301.4	13.9	2869	4	AAH75576
18	235.4	10.9	966	5	AAS70406
19	219.4	10.1	331	3	AAAC76200
20	217	10.0	19025	6	AAAC76200
21	206.6	9.5	2168	4	ABL14125
22	203.8	9.4	735	3	AAAC75784
23	165	7.6	165	2	AAZ14379

24	150	6.9	933	5	AAS72551	Aas72551 DNA encod
25	143.8	6.6	1330	5	AAS70407	Aas70407 DNA encod
26	108.4	5.0	2925	7	ADA52688	Ada52688 Human cod
27	108.4	5.0	5200	6	AAD40740	Aad40740 Human kin
28	108.2	5.0	5234	3	AAC81224	Aac81224 Human his
29	103.6	4.8	599	4	AAL00243	Aal00243 Human rep
30	97	4.5	802	4	AAS03028	Aas03028 Human dia
31	94.6	4.4	439	8	ACH17102	Ach17102 Human adu
32	93.4	4.3	338	4	AAS57421	Aas57421 cDNA #97
33	87	4.0	479	6	ABL81142	Ab181142 Human ova
34	72.4	3.3	500	6	ABS71881	Ab571881 Human GTP
35	71	3.3	167	5	AAS15837	Aas15837 Human pro
36	71	3.3	389	5	AAS72552	Aas72552 DNA encod
37	69.4	3.2	167	6	ABS71866	Ab571866 Human GTP
38	68.2	3.1	789	5	AAS72550	Aas72550 DNA encod
39	68	3.1	2469	6	AAL48837	Aal48837 Human pro
40	68	3.1	2837	7	ABV72511	Abv72511 Nucleotid
41	66.2	3.1	492	3	AAC74860	Aac74860 Human ORF
42	63.6	2.9	2000	7	ADA71938	Ada71938 Rice gene
43	63.2	2.9	114955	2	AAX53491	Aax53491 Human ade
44	62.4	2.9	507	4	AAL34939	Aal34939 Human mus
45	62.4	2.9	507	7	ABX57927	Abx57927 cDNA enco

ALIGNMENTS

RESULT 1

AAL48837

ID AAL48837 standard; cDNA; 2469 BP.

AC AAL48837;

XX 24-OCT-2002 (first entry)

DT Human protein kinase N family kinase coding sequence.
DE Human; protein kinase N; kinase; PKN; eye retinoblastoma; bocio tumour;
KW placenta choriocarcinoma; germ cell; leukocyte; uterus tumour; cancer;
KW pre-B cell acute lymphoblastic leukaemia; Wilm's tumour; enzyme;
KW brain anaplastic oligodendroma; uterus endometrial adenocarcinoma;
KW cytostatic; gene therapy; chromosome 8; gene; ss.

OS Homo sapiens.

PH Key

FT 5'UTR

FT CDS

FT 3'UTR

FT WO200261062-A2.

FT 08-AUG-2002.

FT 29-JAN-2002; 2002WO-US002152.

FT 01-FEB-2001; 2001US-00773371.

FT 07-MAY-2001; 2001US-00849334.

FT (PEKE) PE CORP NY.

FT Rusch D, Ketchum KA, Di Francesco V, Beasley EM;

FT WPI; 2002-608516/65.

FT P-PSDB; AAO18602.

FT New human kinase peptide and nucleic acid molecule, useful for treating

FT disorders associated with abnormal expression of kinase protein, e.g.

FT retinoblastoma, Wilm's tumor, in drug screening assays and

pharmacogenomic analysis.

Claim 4; Fig 1; 76pp; English.

The present invention provides the protein and coding sequences of a novel human protein kinase N family protein. The sequences can be used in the treatment of disorders associated with the absence of, inappropriate, or unwanted expression of the protein, e.g. eye retinoblastoma, Wilms' tumour, placenta choriocarcinomas, boccio tumours, pre-B cell acute lymphoblastic leukaemias, uterine tumours, brain anaplastic oligodendromas and uterine endometrial adenocarcinomas. The present sequence is the cDNA of the invention. The gene is found on human chromosome 8.

Sequence 2469 BP; 461 A; 812 C; 821 G; 375 T; 0 U; 0 Other;

Query Match 100.0%; Score 2169; DB 6; Length 2469;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGATCCTGGAGGAGAGGCCGGACGGCGGGCGCCGGCGAGGAGAGCCCGCGGCTGCAG 60

b
103 ATGATCTGGAGGAGAGGCCGGACGGCGGGCGCCGGCGAGGAGAGCCCGGGCTGCAG 162

61 ATATCTAGGAGAAACCCAGGAAACACGTGTGAGCTCTTTACGGGAAAGACGGGAAGGC 120

163 ATATCTAGGAGAAACCCAGGAAACACGTGTGAGCTCTTTACGGGAAGACGGGAAGGC 222

121 CTGAGAGACGTGTGTGCGTGGAGAGGGTGTCTGGTCCACAGAGGGGAGACCCAGTGCCT 180

223 CTGAGAGACGTGTGTCGCTGGAGAGGGTGTCTGGGTCCACAGAGGGGAAGACCCAGTGCCT 282

181 GTGCACGTTGGCCCCATGAATCCGCAGCTTCATGCAGTGGGCTGTGACTCCCTGACGCAG 240

283 GTGCACGTTGGCCCATGAATCCGCAGCTTCATGCAGTGGCTGTGACTCCCTGACGCAG 342

241 ATCCAGTGGGCCAGCTGCAGAGCCGCAGGGCCAGATTACCCAGCAGATTGACAAGGAG 300

343 ATCCAGTGGGCCAGCTGCAGAGCCGCAGGGCCAGATTCAACAGCAGATTGACAAGGAG 402

301 CTGCAGATGCGGACGGGGCGCTGAGAACCTCTACAGAGCCACCAGCAACAACCGGTGAGA 360

403 CTGCAGATGCGGACGGGCGCTGAGAACCTCTACAGAGCCACCAGCAACAACCGGTGAGA 462

361 GAGACGGTCGCCCTGGAGCTGAGCTACGTCAACTCCAACCTGCAGCTGCTGAAGGAGGAG 420

463 GAGACGGTCGCCCTGGAGCTGAGCTACGTCAACTCCAACCTGCAGCTGCTGAAGGAGGAG 522

421 CTGGAGGAGCTCAGCGGTGGCGTGGACCCCTGGCCGGCATGGGAGCGAAGCTGTCACTGTC 480

523 CTGGAGGAGCTCAGCGGTGGCGTGGACCTGGCCGCATGGGAGCGAAGCTGTCACTGTC 582

481 CCCATGATCCCCCTGGGCCCTGAAGGAGACCAAGGAGCTGGACTGGTCTACACCGCTGAAG 540

583 CCCATGATCCCCCTGGGCCCTGAAGGAGACCAAGGAGCTGGACTGGTCTACACCGCTGAAG 642

541 GAGCTGATCTCAGTGCACCTTGGAGAGGACGGCGCCTCCTACGAGGCAGAAATCAGGGAG 600

643 GAGCTGATCTCAGTGCACCTTTGGAGAGGACGGCGCCTCCTACGAGGCAGAAATCAGGGAG 702

601 CTGGAGGCCCTGCGGCAGGCCATGCGGACCCCGCCGGAATGAGTCGGCCTGGAGCTG 660

703 CTGGAGGCCCTGCGGCAGGCCATGCGGACCCCAAGCCGGAATGAGTCGGGCTGGAGCTG 762

661 CTCACAGCCTATTACAACCAAGCTGTGCTTCTGTGATGCGCGCTTCCTCACCCCTGCCAGG 720

763 CTCACAGCCTATTACAACACAGCTGTGCTTCCTGGATGCGGCTTCCTCACCCCTGCCAGG 822

721 AGCCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTACTGGGTCCCGGCCAGCAGCGT 780

823 AGCCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTACTGGGTCCCGGCCAGCAGCGT 882

781 GCCCTGGCCCTTCGAGAAGGCAGCGTTCTCTTCAACATCGGTGCCCTCCACGCAGATT 840

Db	901	CTCCATGCGCCAGCCAGACATGACGCTGGTCCCTCTCTGGCACTGGAGCAGTCA	1981	TTTCAAGACCCCGCATCACTGGGCGAGTCCCGGCCCCCTCTCAACTGGAGCCGAAA	2040
Qy	999	GATGGCCAGGCCCCAGGAATGTGTGTTTGAAGGCTCTTCAACCACTTGCCTTCAATGCCCC	2079	GGCCAGCAGGCGACAGACTGGAGGCTGCCCCAGCCCTGTGCCAGTGAAGCCAGCTCC	2138
Db	961	GATGGCCAGGCCCCAGGAATGTGTGTTTGAAGGCTCTTCAACCACTTGCCTTCAATGCCCC	2041	GGCCAGCAGGCGACAGACTGGAGGCTGCCCCAGCCCTGTGCCAGTGAAGCCAGCTCC	2100
Qy	1059	CAAAGACTGCTGGCCAGCTGGGCTTGGCGCAGAGGCGCGCCAGGTGGCAGCCAGTA	2139	GGCCTCATCTTTGAAGCACCAGGCTGGCG 2169	
Db	1021	CAAAGACTGCTGGCCAGCTGGGCTTGGCGCAGAGGCGCGCCAGGTGGCAGCCAGTA	2101	GGCCTCATCTTTGAAGCACCAGGCTGGCG 2131	
Qy	1119	CAGGCTAGTGACCGGACCATGCCCCAGCCACCCCTCCAGCACTAGCTGCTCTCTCTG	RESULT 3		
Db	1081	CAGGCTAGTGACCGGACCATGCCCCAGCCACCCCTCCAGCACTAGCTGCTCTCTCTG	ACC46698		
Qy	1179	GACTGCCCTGGTGCATGTCAAGGCGGAGTACTTCCGCTCCCTGGCCCACTTACCACTAGC	ID	ACC46698 standard; cDNA; 3213 BP.	
Db	1141	GACTGCCCTGGTGCATGTCAAGGCGGAGTACTTCCGCTCCCTGGCCCACTTACCACTAGC	XX	AC	ACC46698;
Qy	1239	CATGGCCCTTGGCAGCGCTCCCGACGACCGAGGAGAGCTCCCGCACGACGACGAGGT	XX	02-JUN-2003	(first entry)
Db	1201	CATGGCCCTTGGCAGCGCTCCCGACGACCGAGGAGAGCTCCCGCACGACGACGAGGT	XX	Human dithp biochemical pathway protein-encoding cDNA.	
Qy	1299	CTTCTGACGCCCCCACTCTCTTAAGCCCCGAGGCGCTGTGTGTCGCGCAGAGCTGGA	XX	Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;	
Db	1261	CTTCTGACGCCCCCACTCTCTTAAGCCCCGAGGCGCTGTGTGTCGCGCAGAGCTGGA	KW	cancer; cell proliferative disorder; autoimmune disorder;	
Qy	1359	GGAGCGAGGAGCTTGGCAAGGCACACCTGAAGCGTGCCATCTCTGGGGCAGAGGAGGC	KW	inflammatory disorder; infection; hormonal disorder; metabolic disorder;	
Db	1321	GGAGCGCAGCAGCTTGGCAAGGCACACCTGAAGCGTGCCATCTCTGGGGCAGAGGAGGC	KW	neurological disorder; gastrointestinal disorder; transport disorder;	
Qy	1419	GCTGGGCTGACGCCCCCTGTCGCGGCTCTTGGCGAGTGAGCTGCTTGGGGCTGTGAT	KW	connective tissue disorder; drug screening; proteome analysis;	
Db	1381	GCTGGGCTGACGCCCCCTGTCGCGGCTCTTGGCGAGTGAGCTGCTTGGGGCTGTGAT	KW	gene therapy; antisense therapy; genotyping; transgenic animal; knock in;	
Qy	1479	CTCCACAGCGCTGACGCGCTCACTGGCCAAAGTATGCGAGCTGCACCGTGAGATGACTT	KW	disease model; toxicological testing; transcript imaging;	
Db	1441	CTCCACAGCGCTGACGCGCTCACTGGCCAAAGTATGCGAGCTGCACCGTGAGATGACTT	KW	biochemical pathway; gene; ss.	
Qy	1539	CTGTGAGGCTGCCAGGCCCCCGGACATCCAGCCTTAAGCCACAGAGCCAGAGGCGAG	OS	Homo sapiens.	
Db	1501	CTGTGAGGCTGCCAGGCCCCCGGACATCCAGCCTTAAGCCACAGAGCCAGAGGCGAG	XX	WO200297031-A2.	
Qy	1599	GATGCCAGGCTGTCCAGGGGAGGCGCTGACATCTTCATCGGCTGGGGCCCCCTGTC	XX	05-DEC-2002.	
Db	1561	GATGCCAGGCTGTCCAGGGGAGGCGCTGACATCTTCATCGGCTGGGGCCCCCTGTC	XX	27-MAR-2002; 2002WO-US0100056.	
Qy	1659	TGTCTTCTCAGCCAAAGAACCGGTGGCGGCTGGTGGGCCCCCTCCACCTGACCCGAGGAG	XX	28-MAR-2001; 2001US-0279619P.	
Db	1621	TGTCTTCTCAGCCAAAGAACCGGTGGCGGCTGGTGGGCCCCCTCCACCTGACCCGAGGAG	XX	29-MAR-2001; 2001US-0280067P.	
Qy	1719	GGGCGGCTTGGCTCAGCTTGGGGAGATCTGGCTGTCTCTATCGGTGCGGTCAATTC	XX	29-MAR-2001; 2001US-0280068P.	
Db	1681	GGGCGGCTTGGCTCAGCTTGGGGAGATCTGGCTGTCTCTATCGGTGCGGTCAATTC	XX	16-MAY-2001; 2001US-0291280P.	
Qy	1779	AGGAGCCAGCGCGGCGGCTGGCTCAAGAGGGGAGTACATTTGTGTGAGTGAATGG	XX	17-MAY-2001; 2001US-0291829P.	
Db	1741	AGGAGCCAGCGCGGCGGCTGGCTCAAGAGGGGAGTACATTTGTGTGAGTGAATGG	XX	17-MAY-2001; 2001US-0291849P.	
Qy	1839	GCAGCCATGCAAGTGGTGAGACACGCGAGAGTGGTGA CGGAGCTGAAGGCTGGGGAGA	XX	19-JUN-2001; 2001US-0299428P.	
Db	1801	GCAGCCATGCAAGTGGTGAGACACGCGAGAGTGGTGA CGGAGCTGAAGGCTGGGGAGA	XX	20-JUN-2001; 2001US-0299776P.	
Qy	1899	GGCGGCGCCAGCGCTGAGGTGGTGTGGTGTGCCAGCTTAGACTGCTCCAGCTTGGG	XX	20-JUN-2001; 2001US-0300001P.	
Db	1861	GGCGGCGCCAGCGCTGAGGTGGTGTGGTGTGCCAGCTTAGACTGCTCCAGCTTGGG	XX	(INCY-) INCYTE GENOMICS INC.	
Qy	1959	GGACCGCGGCGGCTCTGCGGCCCCAGGGGCTTTTAAGAGCCAGAGGAGGACATGG	XX	Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;	
Db	1921	GGACCGCGGCGGCTCTGCGGCCCCAGGGGCTTTTAAGAGCCAGAGGAGGACATGG	XX	Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Ameshey SR;	
Qy	2019	TTGCAAGACCCCGGATCCACGTGGGCCAGTCCCGGCCCCCTCTCAACTGGAGCCGAAA	XX	Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;	
			XX	Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;	
			XX	Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;	
			XX	WPI; 2003-129518/12.	
			XX	P-PSDB; ABR41761.	
			XX	Novel human diagnostic and therapeutic polypeptide useful for identifying	
			XX	test compound which specifically binds to a polypeptide encoded by human	
			XX	diagnostic and therapeutic polynucleotide, and to induce antibodies.	
			XX	Claim 2; SEQ ID NO 619; 591pp; English.	
			XX	The invention relates to novel human diagnostic and therapeutic	
			XX	polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded	
			XX	proteins (DITHP; ABR41136-ABR41812). The invention also relates to	
			XX	polynucleotide sequences at least 90% identical to the dithp cDNA	
			XX	sequences of the invention; recombinant vectors, host cells and	
			XX	transgenic organisms comprising a dithp nucleic acid sequence; the	
			XX	recombinant production of DITHP proteins; antibodies specific for DITHP	
			XX	proteins; microarrays comprising dithp nucleic acid sequences; methods of	

detecting dithp nucleotide and protein sequences; methods of screening for compounds which specifically bind a DTHP protein; and methods of assessing the toxicity of test compounds using a dithp hybridisation probe. Dithp nucleic acid sequences and DTHP proteins may be used in the diagnosis of a wide variety of conditions including cancer and other cell proliferative disorders; autoimmune or inflammatory disorders; bacterial, viral, fungal or parasitic infections; hormonal disorders; metabolic disorders; neurological disorders; gastrointestinal disorders; transport disorders; and connective tissue disorders. They may also be used to screen for modulators of protein activity or gene expression. DTHP proteins can additionally be used in analysis of the proteome of a tissue or cell type and to induce antibodies. The dithp nucleic acids are additionally useful in somatic or germline gene therapy of the disorders mentioned above, as a source of antisense sequences, as a source of probes and primers, in genotyping and identification of individuals, in the generation of transgenic animal models of human disease or knock in humanised animals, in toxicological testing, and in transcript imaging. The present sequence represents a dithp cDNA encoding a DTHP protein which is involved in a biochemical pathway. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 3213 BP; 538 A; 1114 C; 1029 G; 532 T; 0 U; 0 Other;

Query Match 36.8%; Score 798.8; DB 7; Length 3213;
Best Local Similarity 92.1%; Pred. No. 1.6e-142;
Matches 853; Conservative 0; Mismatches 72; Indels 1; Gaps 1;

1244 CCCTGTGACGGCTCCCGACGACCGAGGAGAGCTCCCGACGACGACGAGGCTCTCC 1303
712 CCCGTGGCGGCACCCCGACGAAAGTGGCTGTGATGAGCCCGACGACCCCTGGCGTTGC 771
1304 TGCAGCCCCCACCCTCTCTTAAGCCCCGAGGCCCTGTGTGCTGCCGAGGAGCTGAGGAGC 1363
772 CCACCTCTCTTGTCCACGCTCCCGAGGCCCGACGCGCCACATGGTGTGCATCCCGAGTGC 831
1364 GCAGCAGCTTGGCAAGGACACCTGAGCGTGCATCTCTGGGCGAGAGAGGCGCTGC 1423
832 CCGCGTGC-AGGCAAGGACACCTGAAGCGTGCATCTCTGGGCGAGGAGGCGCTGC 890
1424 GGCTGCAGCGCTGTGCGCGCTCTCTGCGCGAGGTGGACCTGCTTCCGCGCTGTGATCTCCC 1483
891 GGCTGCAGCGCTGTGCGCGCTCTCTGCGGAGGTGGACCTGCTTCCGCGCTGTGATCTCCC 950
1484 AGAGCTGCAGCGCTCACTGCGCAAGTATGCGAGCTCGACCGTGAGGATCACTTCTGTG 1543
951 AGAGCTGCAGCGCTCACTGCGCAAGTATGCGAGCTCGACCGTGAGGATCACTTCTGTG 1010
1544 AGGCTGCGAGGCGCCGACATCCAGCTTAAGCCACGACGAGCCAGAGCCGAGGATGC 1603
1011 AGGCTGCGAGGCGCCGACATCCAGCTTAAGCCACGAGCCAGAGCCGAGGATGC 1070
1604 CACGCTGTCCCGAGGAGGCGCTGACATCTTCCATCGGCTGGGCGCCCTGTCTGTGT 1663
1071 CACGCTGTCCCGAGGAGGCGCTGACATCTTCCATCGGCTGGGCGCCCTGTCTGTGT 1130
1664 TCTCAGCCAGAACCGGTGGCGCTGTGTGGGCGCGTCCACTCACCAGGAGAGGCGCG 1723
1131 TCTCAGCCAGAACCGGTGGCGCTGTGTGGGCGCGTCCACTCACCAGGAGAGGCGCG 1190
1724 GCTTTGGCTTCAAGCTTGGGAGACTCGCTGTCTCTCATCGCTGCCGTCAATTCAGGGA 1783
1191 GCTTTGGCTTCAAGCTTGGGAGACTCGCTGTCTCTCATCGCTGCCGTCAATTCAGGGA 1250
1784 GCCAGGCGCGCGGCTGGCTTGAAGGAGGCGGACTACATTTGTCTAGTGAATGGCGAGC 1843
1251 GCCAGGCGCGCGGCTGGCTTGAAGGAGGCGGACTACATTTGTCTAGTGAATGGCGAGC 1310
1844 CATCAGGTGTGTGAGACACCGGAGGTGTGTGACGAGCTGAAGGCTGCGGAGAGGCGG 1903
1311 CATCAGGTGTGTGAGACACCGGAGGTGTGTGACGAGCTGAAGGCTGCGGAGAGGCGG 1370

1904 GCGCCAGCCTGCAGGTGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACC 1963
1371 GCGCCAGCCTGCAGGTGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACC 1430
1964 GCGGGCCGCTGCTGCTGCGGCCCCAGGGGCTCTTAAGGAGCCAGAGGAGCATGTGTGCA 2023
1431 GCGGGCCGCTGCTGCTGCGGCCCCAGGGGCTCTTAAGGAGCCAGAGGAGCATGTGTGCA 1490
2024 AGACCCCGGCATCCACGTGGGCGCAGTCCCGGCCCCCTCTCAACTGGAGCCGAAAGGCC 2083
1491 AGACCCCGGCATCCACGTGGGCGCAGTCCCGGCCCCCTCTCAACTGGAGCCGAAAGGCC 1550
2084 AGCAGGCGAAGACTGGAGGCTGCCCCCGAGCCCTGTGCCCCAGTGAAGCCAGCTCCGCCCT 2143
1551 AGCAGGCGAAGACTGGAGGCTGCCCCCGAGCCCTGTGCCCCAGTGAAGCCAGCTCCGCCCT 1610
2144 CATCTCTGAGACCCAGGCTGGCGG 2169
1611 CATCTCTGAGACCCAGGCTGGCGG 1636

RESULT 4

AAS00831
ID AAS00831 standard; cDNA; 1013 BP.

AC AAS00831;

DT 04-JUL-2001 (first entry)

XX Human cDNA clone HBGDH11 encoding cancer related protein 5.

Human; cancer related protein; HBGDH11; food additive; preservative; immunogen; antibody; bone cancer; adrenal cancer; bone marrow cancer; breast cancer; gastrointestinal cancer; liver cancer; lung cancer; urogenital cancer; immune disorder; Addison's disease; allergy; autoimmune haemolytic anaemia; autoimmune thyroiditis; diabetes mellitus; Crohn's disease; multiple sclerosis; rheumatoid arthritis; ulcerative colitis; acquired immunodeficiency syndrome; AIDS; cardiovascular disorder; myocardial ischaemia; wound healing; neurological disorder; Parkinson's disease; Alzheimer's disease; cerebral anoxia; epilepsy; viral infection; bacterial infection; fungal infection; parasitic infection; agonist; antagonist; ss.

Homo sapiens.

Key	Location/Qualifiers
CDS	6..779
FT	/*tag= a
FT	/product= "Cancer related protein 5"
FT	/partial
XX	/note= "No start codon"

WO200118014-A1.

15-MAR-2001.

30-AUG-2000; 2000WO-US023794.

03-SEP-1999; 99US-0152296P.

06-OCT-1999; 99US-0158003P.

(HUMA-) HUMAN GENOME SCI INC.

Roschke V;

WPI; 2001-235186/24.

P-PSDB; AAS00869.

Twenty nine nucleic acid molecules encoding human cancer associated proteins, useful in the prevention, treatment and diagnosis of cancer, immune disorders, cardiovascular disorders and neurological diseases.

Disclosure; Page 372-373; 427pp; English.

XX The sequence encodes a novel Human cancer related protein. The
CC polynucleotides and polypeptides are useful for preventing, treating or
CC ameliorating a medical condition in e.g. humans, mice, rabbits, goats,
CC horses, cats, dogs, chickens or sheep. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities. The polynucleotide are useful for chromosome
CC identification. The nucleic acids, protein, antibodies, agonists and
CC antagonists are useful in the diagnosis, treatment and prevention of
CC cancer (e.g. cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital), immune disorders
CC (e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,
CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
CC sclerosis, rheumatoid arthritis and ulcerative colitis, acquired
CC immunodeficiency syndrome, AIDS), cardiovascular disorders such as
CC myocardial ischaemias, wound healing, neurological diseases (e.g.
CC Parkinson's disease, Alzheimer's disease, cerebral anoxia and epilepsy)
CC and infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Numerous examples of each type of disorder are given in the
CC specification
XX
SQ Sequence 1013 BP; 197 A; 329 C; 336 G; 151 T; 0 U; 0 Other;
Query Match 36.0%; Score 781.4; DB 5; Length 1013;
Best Local Similarity 99.7%; Pred. No. 2.8e-139;
Matches 793; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1375 GGCAGGACACCTGAGCGTGCATCTCTGGGAGGAGGCGCTGCGGCTGCAGGCC 1434
DB 27 GGCAGGACACCTGAGCGTGCATCTCTGGGAGGAGGCGCTGCGGCTGCAGGCC 86
QY 1435 CTGTGCGGCTCTGCGGAGGTGGACCTCTTGGGCTGTGATCTCCAGAGCGTGCAG 1494
DB 87 CTGTGCGGCTCTGCGGAGGTGGACCTCTTGGGCTGTGATCTCCAGAGCGTGCAG 146
QY 1495 CGCTCACTGCGAAGTATGCGAGCTCGACCGTGGAGTACTTCTGTAGGCTGCCGAG 1554
DB 147 CGCTCACTGCGAAGTATGCGAGCTCGACCGTGGAGTACTTCTGTAGGCTGCCGAG 206
QY 1555 GCCCGGACATCCAGCTAGACCCACAGAGCCAGGCGGAGTGCACGCGCTGTC 1614
DB 207 GCCCGGACATCCAGCTAGACCCACAGAGCCAGGCGGAGTGCACGCGCTGTC 266
QY 1615 CAGGGAGGGGCTGACATCTTCATCGCTGGGCGCCCTGTCTGTGTTCTCAGCCAA 1674
DB 267 CAGGGAGGGGCTGACATCTTCATCGCTGGGCGCCCTGTCTGTGTTCTCAGCCAA 326
QY 1675 AACCGGTGGGCGTGTGTGGGCGCCGCTCCACCTGACCCGAGAGGCGGCTTTGGCCTC 1734
DB 327 AACCGGTGGGCGTGTGTGGGCGCCGCTCCACCTGACCCGAGAGGCGGCTTTGGCCTC 386
QY 1735 ACGTTTCGGGGAGACTGCGCTGTCTCATCGCTCCGCTCATTCAGGGAGCCAGCGCG 1794
DB 387 ACGTTTCGGGGAGACTGCGCTGTCTCATCGCTCCGCTCATTCAGGGAGCCAGCGCG 446
QY 1795 GCGCTGCGCTGAGAGGGGCGACTACATTTGTTCAGTGAATGGCGAGCCATGAGGTGG 1854
DB 447 GCGCTGCGCTGAGAGGGGCGACTACATTTGTTCAGTGAATGGCGAGCCATGAGGTGG 506
QY 1855 TGGAGACACGCGGAGGTGGTGACCGAGCTGAAGGCTGCGGAGAGGCGGCGCCAGCGTG 1914
DB 507 TGGAGACACGCGGAGGTGGTGACCGAGCTGAAGGCTGCGGAGAGGCGGCGCCAGCGTG 566
QY 1915 CAGTGTGTGCTGCTGCCCGACTCTAGATCTGCGCCAGCTTGGGGACCGCGCCCGCTC 1974
DB 567 CAGTGTGTGCTGCTGCCCGACTCTAGATCTGCGCCAGCTTGGGGACCGCGCCCGCTC 626
QY 1975 CTGTGGGCGCCAGGGGCTTTAAGGAGCCAGAGGAGCATGTTGCAAGACCCCGGCA 2034
DB 627 CTGTGGGCGCCAGGGGCTTTAAGGAGCCAGAGGAGCATGTTGCAAGACCCCGGCA 686
QY 2035 TCCACGTGGGCGAGTCCCGCGCCCTCTCACTGGAGCGCAAGGCCAGGCGGCAAG 2094

DB 687 TCCACGTGGGCCAGTCCCGGGCCCTCTCTCAACTGGAGCCGAAAGCCAGCGGCAAG 746
QY 2095 ACTGGAGGTCGCCCGCAGCCCTGTGCCCCAGTGAAGCCAGCTCCGCTCATCTTTGAAG 2154
DB 747 ACTGGAGGTCG-CCCAGGCCCTGTGCCCCAGTGAAGCCAGCTCCGCTCATCTTTGAAG 805
QY 2155 CACCCAGGCGTGGCGG 2169
DB 806 CACCCAGGCGTGGCGG 820
RESULT 5
AAC77780
ID AAC77780 standard; cDNA; 1013 BP.
XX AAC77780;
XX 08-FEB-2001 (first entry)
XX Human cancer associated gene sequence SEQ ID NO:174.
XX Human; cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
KW antidiabetic; antiasthmatic; antirheumatic; antibacterial; antiviral;
KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW vasotropic; antipruritic; antiangiogenic; gene therapy; inflammation;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening; ss.
XX Homo sapiens.
OS
XX
XX WO20005350-A1.
XX 21-SEP-2000.
XX 08-MAR-2000; 2000WO-US005882.
XX 12-MAR-1999; 99US-0124270P.
XX WPI; 2000-587533/55.
XX P-PSDB; AAB43571.
XX Novel isolated nucleic acids comprising sequences encoding peptides
XX useful for treating or diagnosing e.g. cancer.
XX Claim 1; Page 752-753; 2352pp; English.
XX AAC77607 to AAC78448 encode the human cancer associated proteins given in
CC AAB43398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnery; immunomodulator;
CC antidiabetic; antiasthmatic; antirheumatic; antibacterial;
CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
CC nootropic; vasotropic; antipruritic; antiangiogenic and antidiabetic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC rejection, allergic reactions, graft versus host disease and organ
CC disorders, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to

QY	459	TGGAGCGAAGCTGTACATGTGTCCTCCATGATCCCTTGGGGCTGAAGAGACCAAGAGCT	518	QY	1533	TGACTTCTGTGAGGCTGCCGAGGCCCGGACATCCAGCTTAAGACCCACGAGACGAGA	1592
Db	309	CACAGAGGAGGCAATTTACGATTCCTCCCTGATTCCTCTTGGCTTGAAGAAACGAAAGACGT	368	Db	1383	TGACCTGTCTGAACCTGTGATCGACGCCGCCAGTGTGTTGCTTAAACTGAGCAAGAGTTGA	1442
QY	519	GGACTGGTCTACACCCCTGAAGAGCTGATCTCAGTGCACATTTTGGAGAGAGCGCGCTC	578	QY	1593	GGCCAGGATGCCACGCTGTCCAGGGGAAGGGCTTGACATCTTCCATCGGCTGGGGCC	1652
Db	369	CGACTTTGGAGTGGTCTCAAGGATTTTATCTTGGAACTTACAGTGAAGATGGCTATT	428	Db	1443	CATTATATTGCCCCAGTTCTCCAAGCTGACAGTCAAGGACTTCTCCAGAACTGGGCC	1502
QY	579	CTACGAGGACAGAAATCAGGAGCTGGAGGCCCTGCCGAGGCCCATCGGACCCCGAGCCG	638	QY	1653	CCTGTCTGTGTTCTCAGCAAGAACCGGTGGCGGCTGGTGGGGCCCGTCCACCTGACCCG	1712
Db	429	ATATGAGATGAATTCAGATCTTATGGATCTGAGACAGCTTGTTCGAGCGCTAGCCG	488	Db	1503	CTTATCTGTGTTTTCCGCTTAAACAGCGGTGGAGCCCTCCCTCGAAGCATCGCTTCACTGC	1562
QY	639	GAATGAGTGGGGCTGTGAGTGTCTACAGCCTATTACAAACAGCTGTGCTTCTGATGC	698	QY	1713	AGGAGAGGGCGCTTTTGGCTTACCGCTTGGGGAGACTCGCCTGTCTCATCGCTGCCGT	1772
Db	489	GGATGAGGCGGGGTGGAATCTGCTGATGACATATCTATCCAGCTGGGCTTTGTGAGAG	548	Db	1563	AGAAGAAGGGACTTGGGGTTTACCTTTGAGAGGAAACGCCCGCTTCAGGTTCACTTCCT	1622
QY	699	GGCTTCTCTCACCCCTGCCAGGAGCCTCGGGCTCTTCTCCACTGGTACGACTCGCTTAC	758	QY	1773	CATTCCAGGAGCACCGCCCGCGGCTGGCTTGAAGGGGGGACTTACATTGTGTCTCAGT	1832
Db	549	TCGATTTCTTCCCGCCACACGCGAGATGGGACTCTCTGTTCACTGGTATGACTCTCTAC	608	Db	1623	GGATCTCTTACTGCTCTGCTCGGTGGCAGAGGCCCGGGAAGGAGATTATATTGTCTCCAT	1682
QY	759	TGGGGTCCCGCCAGCAGCGTGCCTTGGCTTTGAGAGGGCAGCGTTCTTTCACACAT	818	QY	1833	GAATGGCGACCATGACAGGTGGTGGAGACACGCGGAGGTGGTCAAGGAGCTGAAGGCTGC	1892
Db	609	TGGGGTTCGGTCAAGCCAGCAGAACCTGCTGTGGAGAGGCCAGTGTCTGTTCACAC	668	Db	1683	TCAGCTGTGGATTGAAGTGGCTGACGCTGAGTGAGGTTATGAAGCTGCTGAAGAGCTT	1742
QY	819	CGGTGCGCTCCACAGCAGATTGGGGCGCGCAGGACCGTCTCTGACCCAGAGGTGCCG	878	QY	1893	GGGAGAGGGCGGCGCCAGCCTGCAAGTGGTGTGCTGCTG 1932	
Db	669	TGGGGCGCTCTACACCCAGATTGGGACCCGGTGCATCGGACAGCGCTGGGGCTGGA	728	Db	1743	TGGCGAGGACGAGATCGAGATGAAGTCTGTGAGCCTCCTG 1782	
QY	879	CGCGCTATGGAGGCTTCCAGAGGGCGCTGGGGCTTCAAGCTCTGAGGAGAGAACTT	938	RESULT 7			
Db	729	GAGTGCATAGATGCCCTTTCAGAGCGCGAGGGGTTTAAATTTACCTGGAAGACATTT	788	AA515827			
QY	939	CTCCCATGCGCGGACCCAGACATGAGCGTGCCTTCCCTTGGCGCATGGAGAGCTCAT	998	ID	AA515827	standard; cDNA; 3526 BP.	
Db	789	TACCATCTCCAAATACGACATGAGCCCTGCCATGCTCAGCGTGTCTGTCAAATGAT	848	XX	AA515827;		
QY	999	GATGGCCAGGCGCAGGAATGTGTTGAGGGCTCTCACCACTGCTCCATGGGCC	1058	XX	16-JAN-2002 (first entry)		
Db	849	GCTTGCACAAAGCCGAGGAGCGTGTGAGAAATCAGCTTCTCTG-----GATCCG	902	XX	Human ORF for prostate specific protein PSL22.		
QY	1059	CCAGACTGCTGCGCCAGCTGCGCTGGCGCAGAGGCCGCCAGGTTGGCGAGCGAGTA	1118	XX	Human; prostate specific protein; PSL22; prostate cancer;		
Db	903	GAATGAATCTTCATGCTGTGTAAGTGGCTCAGGAGCTGCTAAGTGGGAGAGTCTA	962	KW	benign prostatic hyperplasia; acute prostatitis; testicular cancer;		
QY	1119	CAGGCTAGTGCACCGGACCATGGCCACCGTCCACGACTACGTCGCTGCTCTCTG	1178	KW	cryptorchidism; testicular disorder; proliferative disorder; lymphoma;		
Db	963	CCACAGCTACAGCGAGCCATGAGCCAGCGCGGTGAAGAGAAACATCCCTTACTCTCTG	1022	KW	leukemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;		
QY	1179	GACTGCCCTGGTGCATGTCAAGGCGGAGTACTTCCGCTCCCTGGCCCACTACCGTAGC	1238	XX	liver cancer; lung cancer; cytostatic; ss.		
Db	1023	GGCCAGCTTAGCTGCGTGAAGGCCCAACACTACGCGGCCCTTGGCCCACTTCTACTGC	1082	OS	Homo sapiens.		
QY	1239	CATGGCCCTCTGCGACGCTCCCGACGACCGAGGAGGCTCCCGACGACGAGCAGGT	1298	XX	Key	Location/Qualifiers	
Db	1083	CATCTCTCTATCGACACACAGGTGAAGCCAGGCAGGATCTGGACCAACGAGGAAATG	1142	FT	1..2061	/tag= a	
QY	1299	CTTCTCTGAGC-----CCCCCACTCTTAAGCCCCGAGGCCCTGTGC--TGCCCGAGA	1352	FT	/product= "PSL22"		
Db	1143	CCTGTCCAGCTCTACGACCAATGCCAGAGGGGCTGACACCCCTTGGCCCACTGAAGAA	1202	XX	WO200172962-A2.		
QY	1353	GCTGAGAGAGCGCAGGAGCTTGGCAAGCACACTGAAGCGTGCATCTCGGGCAGGA	1412	XX	04-OCT-2001.		
Db	1203	TGATCAGAGCGCCGACAGCTGGGAAGTCCCACTTGGCAGAGCATGGCTCATCA	1262	XX	23-MAR-2001; 2001WO-US009410.		
QY	1413	GGAGCGCTTGGCGCTGACGCGCTTGGCGGCTCTGCGCGAGGTGGAACCTGCTCGGGC	1472	XX	24-MAR-2000; 2000US-0191929P.		
Db	1263	GGAGTCGGTGGGAGGGGAGCTCTTGCAAGAGCTGCGGAGCAATTGAGTGTCTACAGAA	1322	PA	(SAAT/) SAATCIOGLU F.		
QY	1473	TGTGATCTCCAGAGCTGACGCGCTCACTGCGCAAGTATGCGGAGCTGACCGTGAGGA	1532	XX	Saatcioglu F;		
Db	1323	GCTGTGTGTGCGCAGGAAAGCTCCCGGCTCAGTACGCCCGACGACGAGGAGGA	1382	DR	WPI; 2001-662926/76.		
				DR	P-PSDB; AAU10192.		
				XX	New polynucleotide for the diagnosis, prevention and treatment for		
				PT	prostate and testis disorders, particularly prostate cancer, comprises		
				PT	prostate-specific or testis-specific nucleic acids.		
				XX	Claim 5; Fig 14A; 114pp; English.		

CC The invention relates to substantially pure prostate-specific or testis-
CC specific polypeptides and the nucleic acids encoding them. Also included
CC are vectors and host cells expressing the proteins, a transgenic animal
CC expressing the protein, antibodies against the proteins, probes for
CC detecting the nucleic acids, antisense molecules for the nucleic acids
CC and methods of isolating modulators of the proteins. Compounds that
CC modulate the prostate specific or testis specific polypeptide are useful
CC to diagnose, prevent or treat disorders of the testis or prostate
CC particularly prostate cancer, benign prostatic hyperplasia, acute
CC prostatitis, testicular cancer, cryptorchidism, undescended, retractile,
CC ascending or vanished testis. Other proliferative disorders for which the
CC modulators may be used include lymphoma, leukaemia, melanoma, ovarian
CC cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer.
CC The present sequence encodes a prostate specific protein, PS422
XX
SQ

Sequence 3526 BP; 992 A; 803 C; 836 G; 895 T; 0 U; 0 Other;

Query Match 22.9%; Score 496.4; DB 5; Length 3526;

Best Local Similarity 56.8%; Pred. No. 5e-85;

Matches 977; Conservative 0; Mismatches 731; Indels 12; Gaps 3;

QY	219	GGGCTGTGACTCCCTGACGAGATCCAGTGGCGGCAGCTGCGAGCCGCGAGGCCACAGAT	278
DB	69	GGGCTGTAATCCCTTGCACAAACCGCGCGAGTAAATTGCAAGATCAAGAGCTCTTT	128
QY	279	TCACCAGCAGATTGACAGGAGCTGCAGATGCGGAGCGGCGCTGAGAACCTCTACAGAGC	338
DB	129	GAATCAGCAGATCTCTGAAGCCGTCGGATGAGGACCGGAGCGGAAACCTTCTGAAAGT	188
QY	339	CACCAGCAACACCGGGTGCAGAGACCGTCCCTGGAGCTGAGCTACGTCACCTCCAA	398
DB	189	GGCCACAACTCAAGGTGGGAGCAAGTGGGCTGGAGCTGAGCTTCTGTCACCTCAGA	248
QY	399	CTGAGCTGCTGAAGGAGGAGCTGGAGAGCTCAGCGGTGGCGTGGACCCCTGGCGGCA	458
DB	249	CTGCAGATGCTCAGGAGAGAGCTGGAGGGGCTGAACATCTCGGTGGCGCTATACAGAA	308
QY	459	TGGGAGCGAAGCTGTCAGTGTCCCATGATCCCCCTGGCGCTGAGGAGACCAAGAGCT	518
DB	309	CACAGAGAGGAGATTTACGATTCCTCTGAGTCCCTGAGGAGGAGGAGGAGGAGAGCT	368
QY	519	GGACTGTGTTACACCGCTGGAAGGAGCTGATCTCAGTGCACCTTTGGAGAGGAGCGGCTC	578
DB	369	CGACTTTGACGTGCTCAGGATTTTATCTTGAAATTTACAGTGAAGATGGCTATTT	428
QY	579	CTAGAGCAGAAATCAGGAGAGCTGGAGCGCTGCGCAGGCCATGCGGACCCCGAGCG	638
DB	429	ATATGAAGATGAAATTGAGATCTTATGGAATCTGAGACAAAGCTTCTCGGACCGCTAGCG	488
QY	639	GAATGAGTCGGGCTGGAGCTGCTCAGACCTATTACAAACAGCTGTCTTCTCGGATGC	698
DB	489	GGATGAGCGCGGTGGAGCTGCTGATGACATATCTTATCCAGCTGGGCTTTGTGAGAG	548
QY	699	GGCTTCTCTACCCCTGTCAGGAGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC	758
DB	549	TCGATTTCTTCCGCGCCACACGCGAGATGGGACTCTGTTCACCTTGTATGACTCTCTAC	608
QY	759	TGGGGTCCCGGCCAGCAGCTGCGCTTTCGAGAAGGCGAGCGTCTCTTCAACAT	818
DB	609	TGGGGTTCGGGTTCAGCCAGCAGAACTCTCTCTGAGAGAGCCAGTGTCTCTTCAACAC	668
QY	819	CGGTGCGCTCCACACGAGATTGGGCGCGCAGAGCCGCTCTGACCGAGGAGTGGCGG	878
DB	669	TGGGGCTCTTACACCCAGATTGGGACCCCGTGGCATCGGAGACGCGAGGCTGGGCTGGA	728
QY	879	CCGGCTATGAGGCTTCCAGAGGCGCTGGGCGCTTACGCTCTCTGAGGAGAACTT	938
DB	729	GAGTGCCATAGATGCTTTTACAGAGGCGCAGGGGTTTAAATTTACCTGAAAGACATT	788
QY	939	CTCCCATGCGCGCCAGCCAGACATGAGCGCTGCGTCCCTCTGCGCACTGGAGAGCTCAT	998
DB	789	TACCCATACTCCAAGTTACGACATGAGCCCTGCCATGCTCAGCGTGTCTGTCAAATGAT	848

QY	999	GATGGCCCAAGGCCCAAGGAATGTGTGTTGAGGGCCCTCTCACCACCTGCTCCATGCCCCC	1058
DB	849	GCTTGCAAGCCCAAGAAAGCGTGTGTTGAGAAATCAGCCCTCTCTG-----GGATCG	902
QY	1059	CCAAGACTGCTGGGCCAGCTGCGCTGGCGAGGAGGCGCCAGGTGGGAGCCGAGTA	1118
DB	903	GAATGAATTTCTTTCATGTGTTGAAGTGGCTCAGGAGGCTGTAAAGTGGGAGAGTCTA	962
QY	1119	CAGGCTAGTGACACCGACCATGGCCCGCAGCCACCGTCCACAGACTAGTGCCTGCTCTCG	1178
DB	963	CCAAACAGCTACAGCAGCATAGCAGCGCGCGGTGAAGAGAACATCCCTACTCTCTG	1022
QY	1179	GACTGCGCTGGTGCATGTCAAGCGCAGTACTTCCGCTTCCCTGGGCCCATCAACAGTAGC	1238
DB	1023	GGCCAGCTTAGCTGCTGCTGAAGGCCCAACCATACGCGGCCCTGGCCCACTACTCTCACTGC	1082
QY	1239	CATGGCCCTCTGGAGGCTCCCGAGCGACCGAGGAGAGCTCCCGACGACGAGAGGT	1298
DB	1083	CATCTCTCTCATCGACCAACAGGTGAAGCCAGGACGAGCTCTGGACCAACGAGAGAGTG	1142
QY	1299	CTTCTGCGAGC-----CCGCCACCTCTCTAAGCCCGGAGGCCCTGTGC--TGCCGACGA	1352
DB	1143	CTGTGCCAGCTCTACGACCAATCCAGAGGGGTGACACCTTTGGCCACACTGAAGAA	1202
QY	1353	GCTGAGGAGCGCAGGAGCTTGGCAAGGCACACCTGAAGCGTGCCTCTGGGCGCAGA	1412
DB	1203	TCATCAGCAGCGCCGACAGCTGGGGAAGTCCACACTTGCAGAGCCATGGCTCATCAGA	1262
QY	1413	GGAGCGCTGGGCTGACAGCGCTGTGCGGCTCTCTGCGGAGGTGAGACTGTCTTGGGCG	1472
DB	1263	GGAGTCGTGGGAGGCGAGCTCTTGAAGAAGCTGCGGAGCATTTGAGGTGCTACAGAA	1322
QY	1473	TGTATCTCTCCAGACCGCTGACAGCGCTCACTGCGCAAGTATGCGAGCTCGACCGTGAGA	1532
DB	1323	GTGCTGTGTGCGGCACAGGAACGCTCCCGCTCAGCTACGTAACGCGCAGCAGCAGGAGAGA	1382
QY	1533	TGACTTCTGTGAGGTGCGGAGGCCCGGAATCAGCGCTTAAGACCCACGAGAACCCAGA	1592
DB	1383	TGACTGTGAAACCTGATCGACGCGCCCGCAGTGTGTTGCTTAAACTGAGCAAGAGTTGA	1442
QY	1593	GGCCAGGATGCCAGCTCTCCAGGGGAGCGGCGCTGACATCTTCCATTCGGCTGGGCGC	1652
DB	1443	CATTATATTTGCCCGCAGTTCTCCAAGCTGACAGTCAAGGACTTCTTCCAGAAAGCTGGGCGC	1502
QY	1653	CCTGTCTGTGTTCTCAGCAAGAACCGGTGGCGCTGGTGGGCGCGCTCCACCTGACCCG	1712
DB	1503	CTTATCTGTGTTTGGGCTAACAGCGGTGAGCGCTCTCGAAGCATCCGCTTCACTTGC	1562
QY	1713	AGGAGGCGCGCTTTGCGCTCAAGCTTGGGAGACTGCGCTGTCTCTCATCGCTGCGCT	1772
DB	1563	AGAAAGAGGGGACTTGGGGTTTCACTTGAGAGGGAACGCCCCCGTTCAAGTTCACTTCT	1622
QY	1773	CATTCCAGGAGCCAGCGCGCGCTGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1832
DB	1623	GGATCTTTATGCTCTGCTGCTGGGAGGAGCCCGGAGGAGGAGGAGGAGGAGGAGGAGG	1682
QY	1833	GAATGGGCGAGCCATGAGGAGTGGTGGAGACAGCGCGAGGTGGTGGAGGAGGAGGAGGAGG	1892
DB	1683	TCAGTTGTGATTTAAGTGGCTGACGCTGAGTGAAGTTATATATTTGCTTCCAT	1682
QY	1893	GGGAGAGCGGGGCCAGCCTGCGAGGTGGTGTGCTGCTG	1932
DB	1743	TGGCGAGGACGAGATCGAGATGAAGTGTGCTGAGCCCTCTG	1782

RESULT 8

ABX97183

ID ABX97183 standard; cDNA; 2109 BP.

XX ABX97183;

XX 20-MAY-2003 (first entry)

XX

QY 699 GGGCTTCTCACCCTGCGAGAGCCCTCGGCTCTTCTTCCACTGAGTACGACTGCTTAC 758
Db 559 TCGATTCTTCCCGCCCAACAGCGAGATGGGACTCTGTTTCACTGGTATGACTCTCTCAC 618
QY 759 TGGGGTCCCGGCCAGCAGCGTCCCTGCGCTTCGAGAGAGGGCAGCGTCTCTTCAACAT 818
Db 619 CGGGGTCCGGTCAAGCAGCAACCTGCTGCTGGAGAGCGCAGTGTCTCTGTTCACAC 678
QY 819 CGGTGCTTCCACACAGCAATGGGGCGGCGCCAGGACCGCTCTTGACCGAGGGTGCCTG 878
Db 679 TGGGGCCCTCTACACCCAGATTGGGACCCGGTGGATGCGACGACGAGCGTGGCTGGA 738
QY 879 CGCGCTATGGAGGCTTCCAGAGGCGCGTGGGGCTTCAGCGCTCTGAGGGAGAACTT 938
Db 739 GAGTGCCATAGATGCTTTTACAGAGAGCCGCGGGGTTTAAATTAACCTGAAAGACACAT 798
QY 939 CTCCATGCGCGAGCCAGACATGAGCGTGGCTCCCTCTGCGCACTGGAGCGCTCAT 998
Db 799 TACCCATCTCCAAGTTACGACATGAGCCCTGCCATGCTCAGCGTCTCGTCAAAATGAT 858
QY 999 GATGGCCCGAGCCAGGAATGTGTGTGAGGCGCTTCAACCACTGCTCCATGCGCCC 1058
Db 859 GCTTGCAAGCCCAAGAAAGCGTGTGGAGAAATCAGCTTCTCTG-----GGATCG 912
QY 1059 CCAAGACTGCTGCGCCAGCTGCGCTGGCTGGCAGAGGCGCCAGGTGGAGCGAGTA 1118
Db 913 GAATGAATCTTCACTGCTGGTGAAGTGGCTCAGGAGGCTGCTAAGGTGGAGAGGTCTA 972
QY 1119 CAGGCTAGTGCACCGGACCATGCGCCAGCCAGCCCGTCCAGCACTAGTGCCTGTCTCTG 1178
Db 973 CCAACAGCTTACAGCGACCATGAGCCAGGCGCGGTGAAGAGAACATCCCTACTCTCTG 1032
QY 1179 GACTGCGCTGTGCGATGTCAAGCGCGAGTATTCCGCTCCCTGGCCCACTACCACTGAG 1238
Db 1033 GCGCAGCTTAGCTGCTGTAAGGCGCCACCACTACGCGGCGCTGGCCCACTACTTCACTGC 1092
QY 1239 CATGCGCTCTGCGAGCGTCCCGAGCGACCGAGGAGGCTCCCGACGACGAGAGGT 1298
Db 1093 CATCTCTCTATGCAACCAAGGTGAAGCCAGGCACTGCTGGACCAACAGGAGAAAGTG 1152
QY 1299 CTCTCTGAGC-----CCCCCACTCTCTTAAGCCCGAGGCGCTGTGC--TGCCGACGA 1352
Db 1153 CTGTGCCAGCTTAGGACCATGCGAGAGGGGTGACACCTTGGCCACACTGAAAGAA 1212
QY 1353 GGTGAGAGCGCAGCGAGCTTGGCAAGGCAACCTGAAGCGTGCCTCTGGGCGACGA 1412
Db 1213 TGATCAGCAGCGCCGACAGCTGGGGAAGTCCCACTTGCGCAGGCAATGGCTCATCAGA 1272
QY 1413 GGAGCGCTGCGGCTGACGCGCTGTCGGGTCTCTGCGGAGGTGGACCTGCTTCGGGC 1472
Db 1273 GAGTCTGTCGGGAGGCAAGCTCTGCAAGAAGCTGCGGAGCAATTGAGGTGCTACAGAA 1332
QY 1473 TGTGATCTCCAGACGCTGCGAGCGCTCACTGCGCAAGTATGGGAGCTCGACCGTGAGGA 1532
Db 1333 GTGTCTGTGTCGCGACAGGAAGCTTCCGGCTCAGTACGCGCCAGCACAGGAGGAGA 1392
QY 1533 TGACTTCTGTAGGCTGCGAGGCGCCGGAATCCAGCTTAAGACCCACCAAGAGCCAGA 1592
Db 1393 TGACCTGTGAACCTGATCGACGCGCCCGCCAGTGTGTTGCTAAAACTGAGCAAGAGTTGA 1452
QY 1593 GGCAGGATGCCAGCGCTTCCAGGGAAGGGGCTTACATCTTCATCGCTGGGGGCC 1652
Db 1453 CATTAATATGCCCCAGTTCTTCAAGCTGACAGTCAAGGACTTCTTCCAGAAAGCTGGGCC 1512
QY 1653 CTGTCTGTGTTCTCAGCAAGAACCGGTGCGGCTGTGGGCGCCGTCACCTGACCCG 1712
Db 1513 CTTATCTGTGTTTTCGGGTAAACAGCGGTGGAAGCGCTCTCGAAGCATCCGCTTCACTGC 1572
QY 1713 AGGAGAGCGCGCTTGGCTCTCAGCTTTCGGGAGACTTCGCTGTCTCTCATCGCTGCCGT 1772
Db 1573 AGAAGAGCGGACTTGGGGTTTCACTTGTAGAGGGGAAGCCCGCCCGTTCAGGTTCACTTCT 1632

QY 1773 CATTCAGGAGCAGGCGCGGCTGCGCTTAAGAGGGCGACTACATTGTGTCACT 1832
Db 1633 GGATCCTTACTGCTCTGCTCGCTCGGTGCGCAGGAGCCCGGAGAGATATATTGTCTCAT 1692
QY 1833 GAATGGGAGCCATGACAGGTGTGTGGAGACACGCGGAGGTGTGACGAGCTGAAGGCTGC 1892
Db 1693 TCAGCTTGTGGATTGTAAAGTGGCTGACGCTGAGTGAGGTATGAAGCTGCTGAAGAGCTT 1752
QY 1893 GCGAGAGGCGGGCGCCAGCGCTGCGAGTGTGTGCTGCTGCTG 1932
Db 1753 TGGCAGGAGCAGATCGATGAAGTCGTGAGCCTCTG 1792

RESULT 9
ABS71856

ID ABS71856 standard; cDNA; 2061 BP.

XX ABS71856;

XX 02-DEC-2002 (first entry)

XX Human GTP-Rho binding protein 2 open reading frame.

XX Human; ss; GTP-Rho binding protein 2; GRBP2; chromosome 19q12; oncogene;
tumour; liposarcoma; ichthyosis congenita III;
benign familial infantile convulsion; gene therapy.

XX Homo sapiens.

XX EP1231216-A2.

XX 14-AUG-2002.

XX 17-JAN-2002; 2002EP-00001026.

XX 30-JAN-2001; 2001WO-US000663.

XX 30-JAN-2001; 2001WO-US000664.

XX 30-JAN-2001; 2001WO-US000665.

XX 30-JAN-2001; 2001WO-US000666.

XX 30-JAN-2001; 2001WO-US000667.

XX 30-JAN-2001; 2001WO-US000668.

XX 30-JAN-2001; 2001WO-US000669.

XX 29-JUN-2001; 2001WO-US000670.

XX 29-JUN-2001; 2001US-00895040.

(AEOM-) ABOMICA INC.

Shannon ME, JI Y;

WPI; 2002-684026/74.

P-PSDB; ABG94709.

Novel GTP-Rho binding protein 2 and nucleic acids encoding the protein,
useful for the manufacture of a medicament for treating a disease
associated with altered expression or activity of human GRBP2 protein.

Claim 1; Page 48-49; 101pp; English.

The invention relates to an isolated GTP-Rho binding protein 2 (GRBP2)
polypeptide or a fragment of at least 6 amino acids or a sequence in
which at least 95% of deviations from GRBP2 sequences are conservative
substitutions. Also included are an isolated nucleic acid (GRBP2 NA)
encoding GRBP2 comprising the full length cDNA or CDS, fragments or
variants, GRBP2 vectors, host cells, antibodies, transgenic non-human
animals modified to contain GRBP2 NA (or unable to express the endogenous
orthologue of GRBP2), diagnosing a disease caused by a mutation in human
GRBP2 or altered expression of GRBP2, anti-agonists of GRBP2, GRBP2
microarrays, fusion proteins and screening for agents that modulate the
expression of GRBP2 NA. GRBP2 is useful for identifying binding partners
of GRBP2. GRBP2, GRBP2 NA and Ab are useful in therapy and in the
manufacture of a medicament for the treatment or prevention of a disorder
associated with increased or decreased expression or activity of human
GRBP2 (e.g. tumours, liposarcoma, ichthyosis congenita III and benign

CC familial infantile convulsion, all associated with the chromosomal
CC location of GRBP2, 19q12). GRBP2 is useful as a standard in immunoassay
CC specific for the proteins, to be used in a therapeutic agent, as
CC vaccines, to be and as antigens (e.g. for epitope mapping) or immunogens
CC (e.g. for raising antibodies. GRBP2 NA is useful as hybridisation probes,
CC to prime synthesis of nucleic acids, to prime first strand cDNA sequence
CC on an mRNA template, and to drive in vivo expression of the proteins. The
CC vector is useful for shuttling GRBP2 NA between host cells derived from
CC disparate organisms, for inserting GRBP2 NA into host cell chromosome,
CC for expressing sense or antisense RNA transcripts of GRBP2 NA in vitro or
CC within a host cell, and for expressing GRBP2 alone or as fusions to
CC heterologous polypeptides. The antibody is useful as an analytical
CC reagent for detection and quantification of GRBP2 and as an immuno
CC therapeutic agent and is useful for flow cytometric detection, for
CC scanning laser cytometric detection, or for fluorescent immunoassay. The
CC present sequence is a GRBP2 cDNA sequence
XX
SQ Sequence 2061 BP; 505 A; 559 C; 574 G; 423 T; 0 U; 0 Other;
Query Match 22.7%; Score 493.2; DB 6; Length 2061;
Best Local Similarity 56.7%; Pred. No. 1.9e-84;
Matches 975; Conservative 0; Mismatches 733; Indels 12; Gaps 3;
QY 219 GGGCTGTGACTCCCTGACGAGATCCAGTGGCCAGCTGCAGAGCCGAGGGCCACAGT 278
DB 69 GGGCTGTGATTCCTTGCACAAACCGCCGGAGTAAATTGCGAGATCAAGAGCTGCTTT 128
QY 279 TCACGACGATTTGACAAAGGAGCTGCAGATGCGGACGGCGCTGAGAACCTTACAGAGC 338
DB 129 GAATCAGCAGATCCTGAAGCCGCTGCGGATGAGGATCGGAGCGGAAACCTTCTGAAAGT 188
QY 339 CACGACAAACACCGGGTGAGAGAGCGGTGCGCTGAGCTGAGCTAGCTCACTCAA 398
DB 189 GGGCACAAACTCAAAGGTGGGGAGCAAGTGGCGCTGGAGCTGAGCTTTCGTAACACTAGA 248
QY 399 CCTGCAGCTCTGAAGGAGAGCTGGAGGAGCTCAGCGGTGGCTGACCCCTGGCCGCGCA 458
DB 249 CCTGCAGATGCTGAAGGAGAGCTGGAGGGGCTGAACATCTCGGTGGCGCTTATCAGAA 308
QY 459 TGGGAGCGAAGCTGTCACTGTCCCATGTATCCCTCGGCTGGAAGGAGACCAAGGAGCT 518
DB 309 CACAGAGGAGGCAATTTACGATTCCTCTGATTCCTCTTGGCTGAAGGAAACGAAAGACGT 368
QY 519 GGACTGTCTACCGCTGAGGAGCTGATCTCATGTGCATTTTGGAGGAGCGCGCTC 578
DB 369 CGACTTTTCAGTCTGCTCTCAGGATTTATCCTGGAACATTTACAGTGAAGATGGCTATT 428
QY 579 CTACGAGCGAAGATCAGGAGCTGGAGGCGCTCGGCGAGGCCATGCGGACCCCGAGCGG 638
DB 429 ATATGAAGATGAATTTGAGATCTTATGGATCTGAGACAAGCTTGTGCGACGCTAGCG 488
QY 639 GAATGAGTCGGGCTGGAGCTGTCTACAGCTATTATCAACACAGCTGTCTTCCTGGATGC 698
DB 489 GGATGAGCGCGGGTGGAACTGCTGATGACATCTTATCATCAGCTGGGCTTTTTCGAGAG 548
QY 699 GCGCTTCTCACCCCTGCCAGAGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC 758
DB 549 TCGATTTCTTCCGGCCACACGCGAGATGGGACTCTGTTCACCTGGTATGACTCTCTCAC 608
QY 759 TGGGCTCCCGCCCGCAGAGCTGGCCCTTCGAGAAGGGCAGCGTTCCTCTTCAACAT 818
DB 609 CGGGGTTCCGCTCAGCCAGCAGAACCTGCTGCTGGAGAAGGCCAGTGTCTCTTCAACAC 668
QY 819 CGGTGCGCTTCCACACGAGATTTGGGCGCGCCAGACCGCTCTGACCCAGGGTGGCCG 878
DB 669 TGGGCGCTCTTACACCCAGATTTGGGACCCGGTGGCATCGGAGACGCGAGGCTGGGCTGGA 728
QY 879 CCGCGCTATGAGGCTTCCAGAGCGCGCTGGGCGCTTCAGCCCTCCTGAGGGAGACTT 938
DB 729 GAGTGCCCATAGATGCTTTTCAGAGAGCGCGAGGGGTTTAAATTACCTTGAAGACACATT 788
QY 939 CTCCCATCGCGCGAGCCAGACATGAGCGCTGCGTCCCTCTGCGCATCTGGAGCACTCAT 998

DB 789 TACCCATACTCCAAAGTTACGACATAGAGCCCTGCCATGCTCAGCGTGTCTCGTCAAAATGAT 848
QY 999 GATGGCCCHAGGCCACAGGAATGTGTGTTGAGGGCTCTCACCACCTGCTCCATGGCCCC 1058
DB 849 GCTTGCACAGGCCAAAGAAAGCGTGTGAGAAAATCAGCCTTCCTG-----GGATCCG 902
QY 1059 CCAAGACTGCTGGCCAGCTGCGCTGGCGCAGGAGGGCGCCAGGTGCGAGCCGAGTA 1118
DB 903 GAATGAATTTCTCATGCTGGTGAAGTGGCTCAGGAGGCTGCTAAGTGGGAGAGTCTA 962
QY 1119 CAGGCTAGTGACCCGAGCAATATGCCCCAGCCAGCCGTCACGACTAGTGTGCTGTCTCTG 1178
DB 963 CCAACAGCTACACGACGACATGAGCCAGGCGCGGTGAAAGAGAAATCCTCCCTACTCTG 1022
QY 1179 GACTGCGCTGGTGCATGTCAAGCCGAGTACTTCCGCTCCCTGCGCCACTTACACAGTAGC 1238
DB 1023 GGCAGCTTAGCTGCTGAAGGCCCACTAGCGGGGCTGACACCTTGGCCACACTTCTACTGC 1082
QY 1239 CATGGCCCTCTGCGACGGCTTCCCCAGCGACCGAGGAGAGCTCCCCACGACGAGCAGGT 1298
DB 1083 CATCTCTCTCATCGACCAACAGCTGAAGCCAGGACGATCTGGACCCAGGAGAGTG 1142
QY 1299 CTTCTGTCAGC-----CCCCCACTCTTAAAGCCCGAGGCCCTGTGC--TGGCGAGGA 1352
DB 1143 CTTGTCCAGCTCTACGACCACTGCGAGAGGGGCTGACACCTTGGCCACACTTGAAGAA 1202
QY 1353 GCTGGAGGAGCGCAGGAGCTTGGCAAGGCACACCTGAAGCGTGCCTATCTGGGGCAGGA 1412
DB 1203 TGATCAGCAGCGCGCAGAGCTGGGGAAGTCCACTTGGCGAGAGCCATGGCTCATCAGGA 1262
QY 1413 GGAAGGCGCTCGGCTGTCACGCCCTGTGCGCGCTCTTGGCGAGGTGAGACTGTCTCGGGC 1472
DB 1263 GAGTGGTGGGGAGGCAAGCTCTGCAAGAGCTGCGGAGCATTCAGGTGCTTACAGAA 1322
QY 1473 TGTGATCTCCAGACGCTGACGCGCTCATGCGCAAGTATGCGGAGCTCGACCTGTAGGA 1532
DB 1323 GGTGCTGTGTGCGCGCACAGGAACGCTCCCGGCTCAOGTACGCCAGCACAGGAGGAGGA 1382
QY 1533 TGACTTCTGTGAGGCTGCGGAGGCCCGGACATCCAGCCTAAGACCCACAGAGCCAGGA 1592
DB 1383 TGACCTGCTGAACCTGATCGACGCCCTCCAGTGTGTGTTGCTTAAACTGAGCAGAGTTGA 1442
QY 1593 GGCAGGATGCCAGCCTGTCCAGGGGAAGGGGCTGACATCTTCCATTCGCTGGGGCC 1652
DB 1443 CATTTATTTGCCCCAGTTCTCAAGCTGACAGTCAAGCTTCTTCCAGAGAGCTGGGGCC 1502
QY 1653 CTTGCTGTGTTCTCAGCAGAGACCGTGGCGGCTGGTGGGGCCCTCCACCTGACCGC 1712
DB 1503 CTTATCTGTGTTTTCGGCTAACAGCGGTGGAOGCTTCTCTGGAAGCATCCGCTTCTAC 1562
QY 1713 AGGAGGCGGCTTTTGGCTTCCAGCTTGGGGGAGACTTCGCTCTCTCTCATCGCTGCGCT 1772
DB 1563 AGAAGAGGGAGCTTGGGGTTTCACTTGAAGGGAAACGCCCTTCAAGGTTCACTTCT 1622
QY 1773 CATTTCCAGGGAGCCAGCCCGCGCTGGGCTGAAAGAGGGCGACTTACATTGTGTCACT 1832
DB 1623 GATCTCTTACTGCTGCTCGCTGGTGGCAGGAGCCCGGAGAGAGTATATTGTCTCCAT 1682
QY 1833 GAATGGGCGAGCCATGAGGTTGGTGGAGACACCGCGAGTGGTGGAGGCTGAAGGCTGC 1892
DB 1683 TCAGCTGTGATTTGATGATGCTGAGCGCTGAGTGGTGTGAGTGTGATGAGCTGCTGAAGAGCTT 1742
QY 1893 GGGAGAGCGGGCGCCAGCCCTGCGAGGTGGTGTGCTGCTGCTG 1932
DB 1743 TGGCGAGGACGAGATCGAGATGAAGTCGTGAGCCTCTG 1782

RESULT 10

ABS71855

ID ABS71855 standard; cDNA; 3484 BP.

XX

AC ABS71855;

XX

DT	02-DEC-2002	(first entry)	Query Match
XX	Human cDNA encoding GTP-Rho binding protein 2.	22.7%; Score 493.2; DB 6; Length 3484;	Best Local Similarity 56.7%; Pred. No. 2e-84;
DE		Matches 975; Conservative 0; Mismatches 733; Indels 12; Gaps 3;	
XX	Human; ss; GTP-Rho binding protein 2; GRBP2; chromosome 19q12; oncogene;		
KW	tumour; liposarcoma; ichthyosis congenita III;		
KW	benign familial infantile convulsion; gene therapy.		
XX			
OS	Homo sapiens.		
XX			
PN	EP1231216-A2.		
XX			
PD	14-AUG-2002.		
XX			
PF	17-JAN-2002; 2002EP-00001026.		
XX			
PR	30-JAN-2001; 2001WO-US000663.		
PR	30-JAN-2001; 2001WO-US000664.		
PR	30-JAN-2001; 2001WO-US000665.		
PR	30-JAN-2001; 2001WO-US000666.		
PR	30-JAN-2001; 2001WO-US000667.		
PR	30-JAN-2001; 2001WO-US000668.		
PR	30-JAN-2001; 2001WO-US000669.		
PR	30-JAN-2001; 2001WO-US000670.		
PR	29-JUN-2001; 2001US-00895040.		
XX			
PA	(AEOM-) AEOMICA INC.		
XX			
PI	Shannon ME, JI Y;		
XX			
DR	WPI; 2002-684026/74.		
DR	P-PSDB; ABG94709.		
XX			
PT	Novel GTP-Rho binding protein 2 and nucleic acids encoding the protein.		
PT	useful for the manufacture of a medicament for treating a disease		
PT	associated with altered expression or activity of human GRBP2 protein.		
PS	Claim 1; Fig 3; 101pp; English.		
XX			
CC	The invention relates to an isolated GTP-Rho binding protein 2 (GRBP2)		
CC	polypeptide or a fragment of at least 6 amino acids or a sequence in		
CC	which at least 95% of deviations from GRBP2 sequences are conservative		
CC	substitutions. Also included are an isolated nucleic acid (GRBP2 NA)		
CC	encoding GRBP2 comprising the full length cDNA or cDS, fragments or		
CC	variants, GRBP2 vectors, host cells, antibodies, transgenic non-human		
CC	animals modified to contain GRBP2 NA (or unable to express the endogenous		
CC	orthologue of GRBP2), diagnosing a disease caused by a mutation in human		
CC	GRBP2 or altered expression of GRBP2, ant-agonists of GRBP2, GRBP2		
CC	microarrays, fusion proteins and screening for agents that modulate the		
CC	expression of GRBP2 NA. GRBP2 is useful for identifying binding partners		
CC	of GRBP2. GRBP2 NA and Ab are useful in therapy and in the		
CC	manufacture of a medicament for the treatment or prevention of a disorder		
CC	associated with increased or decreased expression or activity of human		
CC	GRBP2 (e.g. tumours, liposarcoma, ichthyosis congenita III and benign		
CC	familial infantile convulsion, all associated with the chromosomal		
CC	location of GRBP2, 19q12). GRBP2 is useful as a standard in immunoassay		
CC	specific for the proteins, to be used in a therapeutic agent, as		
CC	vaccines, to be and as antigens (e.g. for epitope mapping) or immunogens		
CC	(e.g. for raising antibodies. GRBP2 NA is useful as hybridisation probes,		
CC	to prime synthesis of nucleic acids, to prime first strand cDNA sequence,		
CC	on an mRNA template, and to drive in vivo expression of the proteins. The		
CC	vector is useful for shuttling GRBP2 NA between host cells derived from		
CC	disparate organisms, for inserting GRBP2 NA into host cell chromosome,		
CC	for expressing sense or antisense RNA transcripts of GRBP2 NA in vitro or		
CC	within a host cell, and for expressing GRBP2 alone or as fusions to		
CC	heterologous polypeptides. The antibody is useful as an analytical		
CC	reagent for detection and quantification of GRBP2 and as an immuno		
CC	therapeutic agent and is useful for flow cytometric detection, for		
CC	scanning laser cytometric detection, or for fluorescent immunoassay. The		
CC	present sequence is a GRBP2 cDNA sequence		
XX			
SQ	Sequence 3484 BP; 932 A; 814 C; 842 G; 896 T; 0 U; 0 Other;		

Db 1103 CATCTCTCTCATGACCAACAGGTGAAGCAGGACGATCTGAGACACACAGAGAGATG 1162
QY 1299 CTTCTCTGACG-----CCCCACCTCTCTAAGCCCGAGGCCCTGTGC--TGCCGACGGA 1352
Db 1163 CTTCTCTGACGCTTAGACCAATGACAGAGGGCTGACACCCCTTGCCACACCTGAAGA 1222
QY 1353 GTTGGAGGACGACGAGCTTGGCAAGGACACCTGAGAGCGTGCATCTCTGGGACGGA 1412
Db 1223 TGATCAGCAGCGGACAGCTGGGGAAGTCCCACTTGGCAGAGCCATGGCTCATCAGCA 1282
QY 1413 GAGGCGCTGCGCTGACGCGCTGTGCGCGTCTCTGCGAGGTGAGACCTGCTCGGGC 1472
Db 1283 GGAGTCGGTGGGAGGAGACCTCTGCAAGAGCTGCGGAGCATTGAGGTGTACAGAA 1342
QY 1473 TGTGATCTCCAGACGCTGACGCGCTCACTGGCCAAAGTATGCGAGCTGCACCTGAGGA 1532
Db 1343 GTTGTCTGTGCGCACAGGAACGCTCCCGCTCAGTACGCCAGCACAGGAGGAGGA 1402
QY 1533 TGAATCTGTGAGGCTGCGAGGCGCCCGGACATCCAGCCTTAAGACCCACAGAGCCAGA 1592
Db 1403 TGAACCTGCTGAACCTGATCGACGCCGCCAGTGTGTGCTTAAACTGAGCAAGAGGTGA 1462
QY 1593 GGCAGAGTCCAGCGCTGTCCAGGAGGAGGCGCTGACATCTTCCATCGCTGGGCGC 1652
Db 1463 CATTAATATGCCAGTCTTCCAGCTGACAGTACAGGACTTCTTCCAGAGCTGGGCC 1522
QY 1653 CTTCTGTCTGTCTTCTCAGCAAGAACCGGTGGCGCTGGTGGGCGCCCTCCACCTGACCG 1712
Db 1523 CTTATCTGTCTTTCGCTAACAAGCGTGGAGCGCTCCTCGAAGCATCCGCTTCACTGC 1582
QY 1713 AGAGAGGCGGCTTGGCTCTCAGCTTGGGAGACTCGCTGTCTCATCGCTGCGCT 1772
Db 1583 AGAAGAGGGGACTTGGGGTTCACCTTGAGAGGGAACGCCCGGTTTCAAGTTCATCTCT 1642
QY 1773 CATTCAGGAGGACGAGCGCGCGCTGGCTTGAAGAGGCGGCTACATTTGTGTGAGT 1832
Db 1643 GATTCCTTACTGTCTGCTGCTGGTGGAGAGCCGCGAGGAGATTAATTTGTCTCAT 1702
QY 1833 GAATGGGACGACATGACAGGTGGTGGAGACACGCGGAGGTGTGACGAGCTGAAGGCTGC 1892
Db 1703 TCAGTTGTGGATTGTAAGTGGCTGACGCTGAGTGAGGTTATGAAGCTGCTGAAGAGCTT 1762
QY 1893 GGGAGAGCGGCGCGAGCTGACGCTGAGTGTGTGCTGCTG 1932
Db 1763 TGGCAGGACGAGATCGAGATGAAGTGTGAGCTCTCTG 1802

RESULT 11

ID AAF58361 standard; cDNA; 3019 BP.

AAAF58361;

19-APR-2001 (first entry)

Human GTP-binding associated protein #61 coding sequence.

Human; guanosine triphosphate binding associated protein; GTP; GBAP;
inflammation; AIDS; Addison's disease; anaemia; arteriosclerosis; asthma;
autoimmune disorder; hepatitis; multiple sclerosis; cancer; diabetes;
osteoporosis; psoriasis; ss.

OS Homo sapiens.

XX WO200105970-A2.

XX 25-JAN-2001.

XX 19-JUL-2000; 2000WO-US019698.

XX 19-JUL-1999; 99US-0144595P.

XX 23-AUG-1999; 99US-0150460P.

PR 15-OCT-1999; 99US-0159849P.
XX (INCY-) INCYTE GENOMICS INC.
XX Yue H, Tang YT, Bandman O, Hillman JL, Lal P, Au-Young J;
PI Reddy R, Yang J, Baughn MR, Lu DAM, Azimzai Y, Patterson C;
XX WPI; 2001-091972/10.
DR P-PSDB; AAB68561.
XX New guanosine triphosphate-binding associated proteins (GBAP) and their
PT encoding nucleic acids, useful for treating and/or diagnosing diseases
XX associated with GBAP expression, such as cancer, diabetes and asthma.
PS Claim 5; Page 229-230; 233pp; English.
XX The present invention relates to novel human guanosine triphosphate (GTP)
CC -binding associated proteins (GBAPs; AAB68561-AAB68566) and their coding
CC sequences (AAF58361-AAF58366). The proteins and coding sequences of the
CC present invention are useful for treating a variety of disorders
CC including inflammation, AIDS, Addison's disease, anaemia,
CC arteriosclerosis, asthma, autoimmune disorders, Grave's disease,
CC hepatitis, multiple sclerosis, cancer, diabetes, osteoporosis and
XX psoriasis
SQ Sequence 3019 BP; 800 A; 708 C; 762 G; 748 T; 0 U; 1 Other;

Query Match 22.5%; Score 487.8; DB 4; Length 3019;
Best Local Similarity 56.9%; Pred. No. 2.1e-83;
Matches 957; Conservative 0; Mismatches 713; Indels 12; Gaps 3;

QY 257 TGCAGAGCCGACGAGCCAGATTACACAGCAGATTCACAGAGGAGCTGCAGATCGGACCG 316
Db 190 TGCAGATCAAGAGCTGCTTTGATCAGCAGATCCTGAAAGCGTGGGATGAGGACCG 249
QY 317 GCCTGAGAACCTTCTACAGAGCCACAGCAACAAACCGGTGTAGAGAGACGGTGCCTTGG 376
Db 250 GAGCGGAAAACCTTCTGAAAGTGGCCACAAACTCAAAGGTGCGGAGCAAGTGCCTGG 309
QY 377 AGCTGAGTACGTCACTCCAACTGCTGCTGAGAGGAGCTGCAGAGCTGCAGCG 436
Db 310 AGCTGAGTCTGCTCACTCAGACCTGAGATGCTCAAGAGAGAGCTGAGAGGCTGACAA 369
QY 437 GTGCGGTGAGACCTTGGCGGCGCATGGGAGCGAAGCTGTCACTGTCCCCTATGCCCTGG 496
Db 370 TCTCGTGGGCTCTATCAGAACACAGAGGAGGCAATTACGATTCCTCTGATCTCTTG 429
QY 497 GCCTGAAAGGAGACCAAGAGCTGAGTGTCTACACCGCTGAGAGAGCTGATCTCAGTGC 556
Db 430 GCCTGAAAGGAAAACGAAAGAGCTGCTGCTGCTCAAGGATTTTATCTCTGGAAC 489
QY 557 ACTTTGGAGAGGACGCGCTCTCTACGAGGAGCAAAATCAGGAGCTGCAGGCGCTTGGCG 616
Db 490 ATTACAGTGAAGATGGCTATTATATGAGATGAATTTGAGATCTTATGATCTGAGAC 549
QY 617 AGGCCATCGGACCCCGCAGCGGAAATGAGTGGGCTTGGAGCTGCTCAGAGCTATTACA 676
Db 550 AAGCTTGTGCGACGCTTAGCCGGATGAGCGCGGGTGGAACTGCTGATGACATACATCTCA 609
QY 677 ACCAGCTGTGCTTCTGATGCGCGCTTCTTCCCTGCGGCTGCAGAGCTCGGGCTCTTCT 736
Db 610 TCCAGCTGGGCTTTGTGAGAGTCAATTTCTTCCCGCCACACGCGAGATGGGATCTCTGT 669
QY 737 TCCACTGGTACGCTCGCTTACTGGGGTCCCGGCGCCAGCAGCGTGCCTTGGCTTCGAGA 796
Db 670 TCACCTGGTATGACTCTCTCAACCGGGTTCGCGTCAGCCAGCAGAACCTGTGCTGGAGA 729
QY 797 AGGCAGCGTTCTCTTCAAGATCGGTGCTTCAACAGCAGATTTGGGCGCGCCAGGACC 856
Db 730 AGGCAGGTGCTCTGTTCAAACACTGGGCGCTTCAACCCAGATTTGGGAGCCCGGTGTGATC 789
QY 857 GCTCTGACACGAGGCTGCGCGCGCTATGGAGGCTTCCAGAGGCGCGCTGGGGCT 916

Db 790 GGCAGACGCGCTGGCTGGAGTGGCCATAGATGCTTTTCAGAGCGCGAGGGGTTT 849
QY 917 TCAGCCTCTGAGGAGAACTTCTCCATGCGCGAGCCAGACATGAGCGCTCGCTGCC 976
Db 850 TAAATTACCTGAAAGACATTTACCCATCTCAAAGTTACGACATGAGCCCTGCCATGC 909
QY 977 TCTGGCAGCTGGAGCAGCTCATGATGGCCAGCCAGCAATGTGTGTTGAGGGCTCT 1036
Db 910 TCAGCGTGTCTGTCAAAATGATCTTGCAAGCCCAAGAAAGCGTGTGTTGAGAAATCA 969
QY 1037 CACCACTGCTCCATGAGCCGCCCAAGACTGCTGCGCCAGCTGCGCTGGCGCAGGAG 1096
Db 970 GCCTTCTCTG-----GATCCNCAATGAATCTTCATGCTGTGAAGTGGCTCAGAGG 1023
QY 1097 CGCCAGGTGGAGCGGAGTACAGGCTAGTGCACCGGACCAATGGCCACGCCACCGCTCC 1156
Db 1024 CTGCTAAGGTGGAGAGGTCTACCAACAGCTACACGACCATGAGCCAGGCGCGGTGA 1083
QY 1157 ACGACTACGTGCTGTCTCTGCACTGCTGCTGTGTGATGTCAGGCGGAGTACTTCGCT 1216
Db 1084 AAGAGAACTCCCTTACTCTCTGGGCCAGCTTAGCTGCGTGAAGGCCACCACTACGCG 1143
QY 1217 CCTGGCCCACTACCACTAGCTAGCCCTCTGCGACGCTCCCGAGCGACCGAGGGAG 1276
Db 1144 CCTGGCCCACTACTTCACTGCACTCTCTCATCGACCAACAGGTGAAGCCAGGCACGG 1203
QY 1277 AGTCCCCAGCAGCAGAGCTTCTCTGAGC-----CCCCACCTCTCTTAAGCCCGA 1332
Db 1204 ATCTGACCAACAGGAGAGTGTCTGCTCCAGCTCTACGACCATGTCAGAGGGGCTGA 1263
QY 1333 GGCCTGTGC--TGCGCGAGCTGGAGGCGGAGGCGCTGCGGTGACGCGCTGCTCTGC 1450
Db 1264 CACCTTGGCCCACTGAAGAATGATCAGAGCGCGCAGCAGCTGGGGAACTCCCACTGC 1323
QY 1391 AGCGTGCCATCTTGGGCGAGGAGGCGCTGCGGTGACGCGCTGTGCGCGCTCTGC 1450
Db 1324 GCAGAGCCATGCTCATCAGGAGTCTGCTGGGAGCGAGCTCTGCAAGAGCTGC 1383
QY 1451 GCGAGTGACCTGCTTGGGCTGTGATCTCCAGACGCTGACGCGCTCACTGGCCAAGT 1510
Db 1384 GGACGATTGAGGTGTACAGAGGTGTGTGTGCGCGCACAGGAACGCTCCCGCTCACT 1443
QY 1511 ATGCGGAGCTCCACCGTGAAGTACTTCTGTGAGGCTGCGGAGCGCGGACATCCAG 1570
Db 1444 AGCCCGAGCACAGGAGGAGTACTCTGTAACCTGATCAGCGCCCCCAGTGTGTG 1503
QY 1571 CTAAAGCCCAAGAGCCAGAGCGCAGATGCCAGCTGTCCAGGGGAGGGGCGCTG 1630
Db 1504 CTAAAGCTGAGCAAGAGGTGTGACATTATATGCCCCAGTTCTCCAAGCTGACGTACGG 1563
QY 1631 ACATCTTCATCGGCTGGGCGCCCTGTCTGTGTTCTCAGCCAAGAACCGGTGGCGGCTGG 1690
Db 1564 ACTTCTTCAAGAGTGGGCGCCCTTATCTGTGTTTTCGGCTAACAGCGGTGGAGCGCTC 1623
QY 1691 TGGGCGCGTCCACTGACCCAGAGAGGCGGCTTGGCCCTCAGCTTCCGGGAGACT 1750
Db 1624 CTCGAGATCCGCTTCACTGCAAGAAGGGGAACTTGGGGTTACCTTGAAGAGGAAG 1683
QY 1751 CGCCTGTCTCATCGCTGCGCTCATTCAGGAGGAGCGAGCGCGCGGCTGCGCTGAAGG 1810
Db 1684 CCCCCTTTCAGTTCCTGCTGGATCTTACTGTCTGCTGCTGCTGGCAGAGGCCCGG 1743
QY 1811 AGGGGACTACATTGTGTGTCAGTGAATGGGCGACCATGAGGTGGTGGAGACACGCGGAG 1870
Db 1744 AAGGAGATTATATTGCTCTCATTGCTGTTGATTTGAATGGCTGACGCTGAGTGAGG 1803
QY 1871 TGGTCAAGGCTGAAGGCTGCGGAGGAGCGCGGCGCGCAGCTGCAAGTGTGCTGCTGC 1930
Db 1804 TTATGAAGCTGCTGAAGAGCTTTGGCGAGGACGAGATCGAGATGAAGTGAAGCTGCC 1863
QY 1931 TG 1932
Db 1864 TG 1865

RESULT 12
ABX97182
ID ABX97182 standard; cDNA; 2310 BP.
XX AC ABX97182;
XX DT 20-MAY-2003 (first entry)
XX DE Human NOV126a cDNA.
XX KW NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
XX KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
XX KW human; gene; ss.
XX OS Homo sapiens.
XX PN W0200272757-A2.
XX PD 19-SEP-2002.
XX PF 08-MAR-2002; 2002WO-US006908.
XX PR 08-MAR-2001; 2001US-0274101P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 08-MAR-2001; 2001US-0274322P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-027676P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0277338P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.

PR 27-SEP-2001; 2001US-0325681P.
 PR 18-OCT-2001; 2001US-0330380P.
 PR 31-OCT-2001; 2001US-0335301P.
 PR 14-NOV-2001; 2001US-0332172P.
 PR 14-NOV-2001; 2001US-0332271P.
 PR 14-NOV-2001; 2001US-0332272P.
 PR 14-NOV-2001; 2001US-0333184P.
 PR 14-NOV-2001; 2001US-0333272P.
 PR 21-NOV-2001; 2001US-0332094P.
 PR 03-DEC-2001; 2001US-0337426P.
 PR 03-DEC-2001; 2001US-0338092P.
 PR 04-DEC-2001; 2001US-0337185P.
 PR 03-JAN-2002; 2002US-0345705P.
 PR 07-MAR-2002; 2002US-00092900.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 XX Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CE, Li L;
 PI Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
 PI Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchervet V;
 PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
 PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
 PI Lepley DM, Rieger DK;
 XX
 XX WPI; 2002-723332/78.
 DR P-PSDB; ABU65215.
 DR
 XX NOVX polypeptides and polynucleotides, useful for preventing or treating
 PT a disorder associated with aberrant NOVX expression or activity e.g.,
 PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
 PT asthma.
 XX
 PS Claim 13; Page 470; 1103pp; English.
 XX
 CC This invention describes novel human NOVX polypeptides which have
 CC cytosolic, cardiac, antiarteriosclerotic, antiasthmatic and hypotensive
 CC activity. Pharmaceutical compositions comprising the NOVX proteins or
 CC nucleic acid molecules or NOVX antibodies are useful for preventing or
 CC treating a disorder associated with aberrant NOVX expression or activity
 CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
 CC asthma. The products of the invention can be used for gene therapy or in
 CC a vaccine. ABX97008-ABX97185 are cDNA fragments amplified and isolated by
 CC the PCR primers and probes represented in ABX13460-ABX13462 and ABX97186-
 CC ABX97593. ABX97008-ABX97185 encode the NOVX proteins described in
 CC ABU65041-ABU65218
 XX
 SQ Sequence 2310 BP; 569 A; 622 C; 628 G; 491 T; 0 U; 0 Other;
 Query Match 21.4%; Score 463.6; DB 6; Length 2310;
 Best Local Similarity 56.7%; Pred. No. 8.1e-79;
 Matches 977; Conservative 0; Mismatches 729; Indels 16; Gaps 6;
 QY 219 GGGCTGTGACTCCCTGACGAGATCCAGTGGCGGAGCTGAGAGCCGAGGCCAGAT 278
 DB 114 GGGCTGTAAATCCCTTGACAAACCGGCGGAGTAAATTCAGAAATCAAGAGCTGCTTT 173
 QY 279 TCACACAGAGATTGACAGAGAGCTGAGATCGGAGCGGCGCTGAGAACCTCTACAGAGC 338
 DB 174 GAATACAGAGATCTCTGAAGCCGCTGCGGATGAGGACCGGAGAAACCTTCTGAAAGT 233
 QY 339 CACCAGCAACACCGGGTGAGAGAGACGGTCCCTCTGGAGCTGAGCTACGTTCAACTCCAA 398
 DB 234 GGCCACAACACTCAAGGTGCGGAGGAGAGTGGCTGGAGCTGAGCTTCGTTCAACTCAGA 293
 QY 399 CCTGAGCTGTGAGAGGAGAGCTGAGAGAGCTAGCGGTGGCTGGACCCCTGGCCGCCA 458
 DB 294 CCTGAGATGTCTAAGGAGAGAGCTGGAGGGGCTGAACATCTCGGTGGCGGTCTTATCAGAA 353
 QY 459 TGGGAGCGAAGCTGTCATGTCCTCCATGATCCCTGGCTGAGGAGAGACCAAGGAGCT 518
 DB 354 CACAGAGGAGGAGATTGATTCCTCCCTGATTCCTTGGCTTGAAGGAAACGAAAGAGCT 413
 QY 519 GGACTGTGTTACACCGCTGAAGAGAGCTGATCTCAAGTGCATCTTTCAGAGAGGAGCGGCGCTC 578

Db 414 CGACTTTGTCAGTGTCTCTCAAGGATTTTATCTTGAACATTTACAGTGAAGTGGCTATTT 473
 QY 579 CTACGAGGAGAAATCAGGAGCTGAGGAGCTTGGAGCCCTGCGGAGCCCATGGGACCCCGCG 638
 Db 474 ATATGAAGATGAATTCAGATCTTATGGATCTGAGACAGCTTGTGCGGAGCTAGCCG 533
 QY 639 GAATGAGTGGGCTGGAGCTGCTCACAGCTATTACACAGCTGTGCTTCTCTGATGC 698
 Db 534 GGATGAGGCGGGGTGAACTGCTGATGATCAATCTTATCATCAGCTGGCTTTGTGAGAG 593
 QY 699 GCGCTTCTCACCCCTGCGCAGGAGCTTGGGCTCTTTCTTCCATGTGTACGACTGCTTAC 758
 Db 594 TCGATTCTCCCGCCACACGCGAGATGGGACTCTCTGTTACCTGGTATGACTCTCTCAC 653
 QY 759 TGGGGTCCCGCCAGCAGGCTGCTCTGGGCTTGGAGAGGAGGAGCGTTCTCTCAACAT 818
 Db 654 CGGGGTTCGGGTGAGCCAGCAACCTGCTGCTGGAGAGGCGAGTGTCTGTTCACAC 713
 QY 819 CGGTGCGCTCCACACGAGATTTGGGGCGCGCCAGGACCGCTCTCTGACCCAGGAGTGC 878
 Db 714 TGGGGCTCTTACACAGATTTGGGACCGGCTGTGATCGGACGCGAGGCTGGGCTGGA 773
 QY 879 CCGCGCTATGAGGCGCTTCCAGAGGCGCTTGGGCGCTTCCAGCTCTCTGAGGAGAACTT 938
 Db 774 GAGTGCATAGATGCTCTTCCAGAGAGCGCGAGGGGTTTAAATTTACCTGAAAGACACAT 833
 QY 939 CTCCATGCGCGAGCCAGCAGATGAGCGCTGCTGCTCTCTGCGCAGCTGAGCAGGCTCAT 998
 Db 834 TACCCATCTTCCAAAGTTACGACATGAGCCCTGCCATGCTCAGCGTGTCTGTCAAATGAT 893
 QY 999 GATGGCCCGAGCCCGAGGAATGTGTGTTGAGGGGCTCTCACACCTGCTCTCATGGCCCC 1058
 Db 894 GCTTGCACAGCCCAAGAAAGCGTGTGAGAAATCAGCCTTCTCTG-----GGATCCG 947
 QY 1059 CCAAGACTGTGCGCCAGCTGCGCTGCGGCGAGGAGCGCCCGAGGTGGCAGCCGAGTA 1118
 Db 948 GAATGAATTTCTTATGCTGTGTAAGTGTGCTCAGGAGGCTGCTAAGTGTGGAGAGTCTA 1007
 QY 1119 CAGGCTAGTCACCGGACCATGGCCAGCCAGCCGCTCAGAGCTACGCTGTCTCTCTG 1178
 Db 1008 CCAACAGCTACAGCAGCCATGAGCAGGCGCGCTGTAAGAGAAACATCCCTACTCTG 1067
 QY 1179 GACTGCTCTGCTGATGCTCAAGCGCGAGTACTTCCGCTCTCTGCGCCCACTACACGTA 1238
 Db 1068 GGCAGCTTAGCTGCTGTAAGGCCACCACTAGCGGCGCTGCGCCCACTACTTCACTGC 1127
 QY 1239 CATGGCCCTCTGCGACCGGCTCCCGAGCGACGAGGAGAGCTCCCGCAGCAGGAGAGT 1298
 Db 1128 CATCTCTCTCATCCAGCCAGGCTGAAGCCAGGCGAGATCTGGACCAACGAGGAGAGT 1187
 QY 1299 CTTCTCTCAGC-----CCCCCAGCTCTCTTAAGCCCGAGGCGCTGTGC--TGCCGCGAGA 1352
 Db 1188 CCTGTCCAGCTCTACGACACATGCGAGAGGGCTGACACCTTGGCCACACTGAGAA 1247
 QY 1353 GCTGGAGAGCGCAGGAGCTTGGCAAGGACACCTGAAGCGTGCCTCTCTGGGGCAGGA 1412
 Db 1248 TGATCAGCAGCGCGCAGCAGCTGGGGAAGTCCCACTTGGCAGAGCCATGGCTCATCAGA 1307
 QY 1413 GGAGCGCTGGGCTGACGCGCTGTGCGGCTCTGCGGAGGTGGACCTGCTTCCGGC 1472
 Db 1308 GGAGTCGTTGGGAGGCGCAGCTCTGCAAGAACTCTGGAGAGCAATTGAGGTGCTACAGAA 1367
 QY 1473 TGTGATCTCCAGAGCGCTGACAGCGCTCACTGGCCCAAGTATGCGGAGCTCGACCGTGA 1532
 Db 1368 GGTGCTGTGTCGCCGACAGGACCGCTCCCGCTCACTGACCGCCAGCAGGAGGAGGA 1427
 QY 1533 TGACTTCTGTGAGCTGCGGAGGCGCC--GGACATCCAGCTTAAGACCCACCAAGAGCCA 1590
 Db 1428 TGACCTGTGAACCTGATCGACGCGCCCGAGAGTGTGTTGTTAAACTGAGCAAGAGGTT 1487
 QY 1591 GAGCCAGGATGCCACGCTGTCCAGGAGGAGGCGCTGATCTTCCATCGGCTGGG 1650

Db 1075 GAGCCTCTGCAAGAGCTCCG-GNATTGAGTCTTACAGAGGTGTGNGTGGCG-CACA 1132
QY 1491 GCAGCGCTCACTGCGCCAGTATGCGGAGCTCGACCGTGAGAGTACTTCTGTGAGGCTGC 1550
Db 1133 GGAACGCTCCCGCTACGCTACGCGCCAGCACCCAGGAGGATGACCTGCTGAACCTGAT 1192
QY 1551 CGAGGCGCCGACATCCAGCCTTAAGACCCACCAAGCCAGAGCCAGAGTGCACGCCT 1610
Db 1193 CGAGCGCCCGAGTGTGTTGCTTAAACTGAGCAAGAGTTGACATTATATGCGC-CCATT 1251
QY 1611 GTCCGAGGGAAGGGGCTGACATCTTCCATCGGCTGGGCGCCCTGCTGTGTTCTCAGC 1670
Db 1252 CTCGAAGCTGACAGTCAAGGACTTCTTCCGAAGCTGGGCGCCCTATCTGTTGTTTCGGC 1311
QY 1671 CANGAACGGTGGCGGTGTGGGGCGCGCTCCACCTGACCCGAGGAGGGCGGCTTTGG 1730
Db 1312 TAAACAAGCGGTGACGCTCTCTCGAAGCATCCGCTTCACTGCAGAGAGAGGGGACTTGG 1371
QY 1731 CCTCAGCTTCGGGGAGCTCGCTCTGCTCATCGCTGCGCTCATTCAGGAGCCAGGC 1790
Db 1372 GTTCACTTCAGAGGGAACGCCCGCTTCAAGTTCACCTTCTCGATCCTTACTGCTGTC 1431
QY 1791 CGCGCGGCTGGCTGAAGAGGGGCGACTACATTGTGTCAATGATGGGAGCCATGCGAG 1850
Db 1432 CTCGGTGGCAGGAGCCCGGGAAGAGATTATTTGCTCCATTCAGCTTGTGATTTGTA 1491
QY 1851 GTGGTGGACACGCGGAGGTGTGAGCGAGCTGAAGGCTGAAGGCTGCGGAGAGG 1900
Db 1492 GTGGCTGACGCTGAGTGAGGTTATGAAGCTGTGAAGAGCTTGGGCGAGG 1541

RESULT 15

ABK92255
ID ABK92255 standard; DNA; 1671 BP.

XX
AC ABK92255;

XX
DT 15-AUG-2002 (first entry)

XX
DE Prostate cancer-associated DNA sequence #141.

XX
KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;
KW gene therapy; gene; ds.

XX
OS Mammalia.

XX
PN WO200230269-A2.

XX
PD 18-APR-2002.

XX
PF 12-OCT-2001; 2001WO-US032045.

XX
PR 13-OCT-2000; 2000US-00687576.

PR 08-DEC-2000; 2000US-00733288.

PR 08-DEC-2000; 2000US-00733742.

PR 24-JAN-2001; 2001US-0263957P.

PR 16-MAR-2001; 2001US-0276791P.

PR 16-MAR-2001; 2001US-0276888P.

PR 06-APR-2001; 2001US-0281922P.

PR 30-APR-2001; 2001US-0286214P.

PR 04-MAY-2001; 2001US-00847046.

XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX
PI Gieh KC, Mack DH, Wilson KE, Afar D, Hevezi P;

XX
DR WPI; 2002-471335/50.

XX
DR P-PSDB; ABG61936.

XX
PT Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a

PT prostate tissue.

XX
PS Claim 22; Page 417-418; 436pp; English.

XX
CC The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with prostate cancer-
CC associated polynucleotides (designated PC genes) that selectively
CC hybridise to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
CC sequences

XX
SQ Sequence 1671 BP; 402 A; 459 C; 457 G; 353 T; 0 U; 0 Other;

Query Match 17.5%; Score 379.2; DB 6; Length 1671;
Best Local Similarity 56.0%; Pred. No. 8.7e-63;

Matches 783; Conservative 0; Mismatches 603; Indels 12; Gaps 3;

QY 541 GAGCTGATCTCAGTGCACCTTTGAGAGGAGGGGCGCTCTACGAGCGAATACAGGAG 600
Db 1 GATTTTATCTCGAACATTACAGTGAAGATGGCTATTATATGAAGATGAATTCAGAT 60

QY 601 CTGGAGGCGCTCGGCGAGGCCATGCGGACCCCGACCGGAATCAGTCGGCGCTGGAGCTG 660
Db 61 CTTATGGATCTGAGACAGCTTGTGCGACCGCTAGCCGGGATGAGCCGGGGTGGAACTG 120

QY 661 CTCACAGCCTATTACAAACAGCTGTGCTTCTGGATGCGCGCTTCTCAACCTCCAGCG 720
Db 121 CTGATGACATACTTCATCCAGCTGGGCTTTGTGAGAGTCTGATTCTTCGCGCCACAGG 180

QY 721 AGCCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTACTGCGGTCCCGGCCAGAGCT 780
Db 181 CAGATGGGACTCTCTGTTCACCTGGTATGACTCTCTCACGGGGTTCGGGTCCAGCCAG 240

QY 781 GCGCTGGCGCTTCGAGAAGGGCAGCTTCTCTCAACATCGGTGCGCTCCACACAGAGATT 840
Db 241 AACCTGCTGCTGGAGAGGCGAGTGTCTGTGTTCAACATGCGGGGCGCTCTACACCCAGATT 300

QY 841 GGGGCGCGCCAGGACCGCTCTCCGACCGAGGGGTGCGCGCGCTATGAGGCGCTTCCAG 900
Db 301 GGGACCCGCTGTGATCGGCGAGCGAGGCTGGGCTGGAGAGTGCCATAGTGCCTTTTCAG 360

QY 901 AGGGCGCGTGGGCGCTTACGCTCTCTGAGGAGAACTTCTCCATGCGCGCCAGCCAGAC 960
Db 361 AGAGCCGCGAGGGGTTTAAATTTACCTGAAAGACATTTACCCATACTCAAGTTTACGAC 420

QY 961 ATGAGCGCTGCGCTCCCTCTGCGCACTGGAGCAGCTCATGATGGCCCGCCAGCGAATGT 1020
Db 421 ATGAGCCCTGCGCATGCTCAGGTGCTCTCAATGATGCTTGCACAGCCCAAGAGC 480

QY 1021 GTGTTTGAAGGCGCTCTACACCTGCTCCATGGCCCCCGCCAGAGCTGCTGCGGCCAGCTG 1080
Db 481 GTGTTTGAAGAAATCAGCCTTCTCTG-----GGATCCGGATGAATCTTCATGCTGTG 534

QY 1081 CGCCTGGCGCAGGAGGCGCGCCAGTGCGACCGAGTACAGGCTAGTGCAACCGGACCATG 1140
Db 535 AAGGTGGCTCAGGAGGCTGTAAAGTGGGAGAGGTCTACCAACAGCTACACGAGCCCATG 594

QY 1141 GCCCAGGCCACCCGCTCCACGACTAGTGTGCTCTCTCTGGACTGCTCCCTGGTGCATGTCAG 1200
Db 595 AGCCAGCGCGGTGAAGAGAACATCCCTACTCTCTGGGCCAGCTTAGCCTGCGTGAAG 654

QY 1201 GCCGAGTACTTCCGCTCCCTGGGCCACTACACAGTAGCCATGGCCCTCTGCGAGGCTCC 1260
Db 1201 GCGGAGTACTTCCGCTCCCTGGGCCACTACACAGTAGCCATGGCCCTCTGCGAGGCTCC 1260

Db 655 GCCCACCACCTACCGGCGCCCTGGCCCACTTCTCCTGCCATCCTCTCATCGACACCAG 714
Qy 1261 CCAGCGACCGAGGAGAGCTCCACGACGACGAGCGTCTTCTCTGAGC----CCCCAC 1316
Db 715 GTGAAGCCAGGACCGGATCTGGACCAACCGAGAGAGTGCCCTGTCCAGCTCTACGACCAC 774
Qy 1317 CTCTCTTAAGCCCGGAGGCGCTGTGCT--GCCCGAGGAGCTGGAGGCGCAGGAGCTT 1374
Db 775 ATGCCAGAGGGGTGACACCTTTGGCCACACTGAAGAAATGATCAGCAGCGCCGACAGCTG 834
Qy 1375 GGCAAGGCACACCTGAAGCGTGCCATCCTGGGGCAGGAGGAGCGCTGCGGCTGCACGCC 1434
Db 835 GGAAGTCCCACTTGGCAGAGCCATGGCTCATCAAGAGAGTCGGTGCAGGAGGCCAGC 894
Qy 1435 CTGTGCGCGCTCTCGCGAGGCTGAGCTCTCTGGGCTGTGATCTCCAGACGCTGCAG 1494
Db 895 CTCTGCAAGAAGCTGCGGAGCAATTAGGTGCTACAGAAGGTGCTGTGCGCCACAGGAA 954
Qy 1495 CGTCACTGCCCAAGTATGCGGAGCTCGACGCTGAGGATGACTTCTGTGAGGCTGCCGAG 1554
Db 955 CGTCCCGGCTCAGTACGCGCCACGACCGAGGAGGATGACCTGTGAACCTGATCGAC 1014
Qy 1555 GCCCGGACATCCAGCCTAAGACCCACGAGCCAGAGCCAGGATGCCACGCGCTGCC 1614
Db 1015 GCCCCAGTGTGTGTGTAACCTGAGCAAGAGGTTGACATTATTTGCCCCAGTTCTCC 1074
Qy 1615 CAGGGGAAGGGGCTGACATCTTCATCGGCTGGGGCCCTGTCTGTGTGTTCTCAGCCAAG 1674
Db 1075 AAGCTGACAGTCACGGACTTCTCCAGAAAGCTGGGCGCCCTTATCTGTGTTTTCGGCTAAC 1134
Qy 1675 AACGCTGGCGCTGGTGGGCGCCCTCCACTGACCGGAGAGGGCGGCTTTGGGCTC 1734
Db 1135 AAGCGTGGACGCTCTCGAAGCATCCGCTTCACTGCAAGAAAGGGGACTTGGGGTTC 1194
Qy 1735 ACGCTTCGGGAGACTCGCCTGTCTCATCGCTGCCCTCAATTCAGGAGAGCCAGGCGCG 1794
Db 1195 ACCTTGAGAGGGAACGCCCGCTTCAGGTTCACTTCTGGATCCTTACTGCTCTGCCCTCG 1254
Qy 1795 GCGGCTGCGCTGAAGGAGGGCGACTACATTTGTGTGATGAATGGGCGAGCCATGCAAGTGG 1854
Db 1255 GTGGCAGGAGCGCCGGAAGGAGATTATATTGCTCTCCATTCACTTGTGGATTGTAAGTGG 1314
Qy 1855 TGGAGACACGGGAGGTGGTACGAGCTGAAGGCTGGGAGAGGGCGGCGCCAGGCGTG 1914
Db 1315 CTGACGCTGAGTGAGGTTATGAAGCTCTGAAGAGCTTTGGCAGGACGATCGAGATG 1374
Qy 1915 CAGGTGCTGCTGCTG 1932
Db 1375 AAAGTCTGAGCCTCTG 1392

Search completed: July 13, 2004, 02:51:08
Job time : 564.787 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2004, 20:08:35 ; Search time 5564.67 Seconds
(without alignments)
16894.276 Million cell updates/sec

Title: US-10-697-266-1_COPY_103_2271

Perfect score: 2169

Sequence: 1 atgatactcggaggagagcc.....tgaagcaccacgggtggccg 2169

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2169	100.0	2469	6	AR269115	Sequence
2	2169	100.0	2469	6	AX642309	Sequence
3	1951	89.9	2837	6	AX675063	Sequence
4	1951	89.9	3763	9	BC025767	Homo sapi
5	1866	86.0	2088	9	AY082588	Homo sapi
6	1865	86.0	3647	9	AB067516	Homo sapi
7	1102.2	50.8	1932	10	MMU43194	Mus muscu
8	1036.6	47.8	2301	10	BC052010	Mus muscu
9	798.8	36.8	4729	6	BD183400	Novel gen
10	558.8	25.8	3230	4	CFA347749	Canis fam
11	496.4	22.9	3516	9	BC036447	Homo sapi
12	494.8	22.8	2061	9	HSA347750	Homo sapi
13	494.8	22.8	3500	6	AX835273	Sequence
14	494.8	22.8	3500	9	AK098246	Homo sapi
15	487.8	22.5	3019	6	AX077672	Sequence
16	485.6	22.4	3397	9	HSN803254	Sequence
17	477.4	22.0	2233	5	BC044556	Danio rer
18	475.8	21.9	2224	5	AY391456	Homo sapi
19	443.4	20.4	2262	9	AF268032	Homo sapi
20	436	20.1	764	6	BD214873	Novel hum
21	418.2	19.3	2757	6	AX593072	Sequence
22	311	14.3	3798	9	AK126506	Homo sapi
23	272.8	12.6	2403	3	AF132025	Drosophil
24	222.4	10.3	178343	2	AC024049	Homo sapi
25	217	10.0	19025	6	AR269116	Sequence
26	217	10.0	19025	6	AX642311	Sequence
27	217	10.0	179941	9	AF276758	Homo sapi
28	217	10.0	179941	9	AC105118	Homo sapi
29	168.2	7.8	61188	2	AC135076	Homo sapi
30	165.4	7.6	123016	2	AF189001	Homo sapi
31	165.4	7.6	124060	9	AC087793	Homo sapi
32	165	7.6	165	6	BD213706	Novel hum
33	149.6	6.9	2300	6	AX833345	Sequence
34	149.6	6.9	2300	9	AK095001	Homo sapi
35	146.6	6.8	232687	2	AC116393	Mus muscu
36	143.4	6.6	261608	2	AC097158	Rattus no
37	143.4	6.6	272301	2	AC133673	Rattus no
38	127	5.9	3174	2	AC014171	Drosophil
39	127	5.9	185200	3	AC094401	Drosophil
40	127	5.9	191558	3	AC013431	Drosophil
41	127	5.9	303345	3	AE003500	Drosophil
42	110	5.1	5234	9	AB025194	Homo sapi
43	108.4	5.0	2925	6	AX713572	Sequence
44	108.4	5.0	2925	9	AK055157	Homo sapi
45	108.4	5.0	4265	9	AB040904	Homo sapi

ALIGNMENTS

RESULT 1
AR269115
LOCUS AR269115
DEFINITION Sequence 1 from patent US 6500655.
ACCESSION AR269115
VERSION AR269115.1 GI:29699963
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2469)
AUTHORS Rusch,D., Ketchum,K.A., Di Francesco,V. and Beasley,E.M.
TITLE Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof
JOURNAL Patent: US 6500655-A 1 31-DEC-2002;

2469 bp DNA linear PAT 10-APR-2003

[illegible]

Db 1663 GACATCCAGCTTAAGACCCACAGAAAGCCAGAGGCCAGAGTGCACGCCTGTCCACGGGG 1722
QY 1621 AAGGGGCTTACATCTTCATTCGGCTTGGGGCCCTCTGTGTCTCAGCAAGAACCGG 1680
Db 1723 AAGGGGCTTACATCTTCATTCGGCTTGGGGCCCTCTGTGTCTCAGCAAGAACCGG 1782
QY 1681 TGGCGGCTGGTGGGGCCCTGCACCTGACCCGAGGAGAGGGCGGCTTTGGSCCTCAGCTT 1740
Db 1783 TGGCGGCTGGTGGGGCCCTGCACCTGACCCGAGGAGAGGGCGGCTTTGGSCCTCAGCTT 1842
QY 1741 CGGGAGAGCTCGCTGTCTCATCTGCTGCTGCTGCTCAATTCAGGGAGCCAGGCGGGGGCT 1800
Db 1843 CGGGAGAGCTCGCTGTCTCATCTGCTGCTGCTCAATTCAGGGAGCCAGGCGGGGGCT 1902
QY 1801 GGCCTGAAGAGAGGGCGGCTACATTTGTCTAGTGAATGGGAGCCATGCAAGTGGTGAGA 1860
Db 1903 GGCCTGAAGAGAGGGCGGCTACATTTGTCTAGTGAATGGGAGCCATGCAAGTGGTGAGA 1962
QY 1861 CAGCGGAGGTGGTACGGAGCTGAGGCTGCGGGAGAGCGGGCGCCAGCTTGCAGGTG 1920
Db 1963 CAGCGGAGGTGGTACGGAGCTGAGGCTGCGGGAGAGCGGGCGCCAGCTTGCAGGTG 2022
QY 1921 GTGTGCTGTGCTGCCAGCTCTAGACTGCCAGCTTGGGGAGCCGCGGCGGCTGCTGCTG 1980
Db 2023 GTGTGCTGTGCTGCCAGCTCTAGACTGCCAGCTTGGGGAGCCGCGGCGGCTGCTGCTG 2082
QY 1981 GGGCCCAAGGGGCTTTAAGAGCCAGAGGAGCATGGTTGCAAGACCCCGGCATCCAGC 2040
Db 2083 GGGCCCAAGGGGCTTTAAGAGCCAGAGGAGCATGGTTGCAAGACCCCGGCATCCAGC 2142
QY 2041 TGGGCGAGTCCCGGCGGCTTCACTGAGCGCCGAAAGGCCAGAGGCGCAAGACTGGA 2100
Db 2143 TGGGCGAGTCCCGGCGGCTTCACTGAGCGCCGAAAGGCCAGAGGCGCAAGACTGGA 2202
QY 2101 GGTGCGCCCGAGCGCTGTGCGCCAGTGAAGCAGCTCCGCGCTCATCTTGAAGACCCA 2160
Db 2203 GGTGCGCCCGAGCGCTGTGCGCCAGTGAAGCAGCTCCGCGCTCATCTTGAAGACCCA 2262
QY 2161 GGGTGGGCG 2169
Db 2263 GGGTGGGCG 2271

RESULT 3
AX675063 2837 bp DNA linear PAT 27-MAR-2003
LOCUS Sequence 11 from Patent WO02077235.
DEFINITION AX675063
ACCESSION AX675063
VERSION AX675063.1 GI:29333362
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Burford,N., Ding,L., Yue,H., Thornton,M., Wallia,N.K., Gandhi,A.R.,
Arvizu,C., Baughn,M.R., Swarnakar,A., Duggan,B.M., Lu,D.A.,
Thangavelu,K., Warren,B.A., Tang,F.I., Khan,F.A., Yao,M.G. and
Emerling,B.M.
TITLE Intracellular signaling molecules
JOURNAL Patent: WO 02077235-A 11 03-OCT-2002;
Incyte Genomics, Inc. (US)
FEATURES Location/Qualifiers
source
1..2837
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 5560161CBI"

ORIGIN
Query Match 89.9%; Score 1951; DB 6; Length 2837;
Best Local Similarity 100.0%; Pred. No. 8.1e-295;
Matches 1951; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 219 GGGCTGTGACTCCCTGACGAGATCCAGTCCGCGCAGCTGCAGAGCCGAGGGCCAGAT 278
Db 181 GGGCTGTGACTCCCTGACGAGATCCAGTCCGCGCAGCTGCAGAGCCGAGGGCCAGAT 240
QY 279 TCACAGCAGATTTGACAAAGAGCTGCAGATGCGGACGGGGCTGAGAACTCTACAGAGC 338
Db 241 TCACAGCAGATTTGACAAAGAGCTGCAGATGCGGACGGGGCTGAGAACTCTACAGAGC 300
QY 339 CACACGACACACCGGGTGCAGAGAGCGGTGCGCCCTGGAGCTGAGCTACGCTCAACTCCAA 398
Db 301 CACACGACACACCGGGTGCAGAGAGCGGTGCGCCCTGGAGCTGAGCTACGCTCAACTCCAA 360
QY 399 CTTGCAAGCTGCTGAAAGAGAGCTGCAGAGAGCTCAGCGGTGGCGTGGAACCTGGCGGCA 458
Db 361 CTTGCAAGCTGCTGAAAGAGAGCTGCAGAGAGCTCAGCGGTGGCGTGGAACCTGGCGGCA 420
QY 459 TGGGAGCGAAGCTGTCTGCTCCCATGATCCCGCTGGGGCTGAAAGAGACCAAGAGCT 518
Db 421 TGGGAGCGAAGCTGTCTGCTCCCATGATCCCGCTGGGGCTGAAAGAGACCAAGAGCT 480
QY 519 GGAAGCTGTACACCGCTGAAGAGAGCTGATCTCAGTGCACCTTTGGAGAGACCGCGCTC 578
Db 481 GGAAGCTGTACACCGCTGAAGAGAGCTGATCTCAGTGCACCTTTGGAGAGACCGCGCTC 540
QY 579 CTACAGAGCGCAAGAAATCAGGAGCTGCAGGCGCTTGCGGCGAGGCCATGCGGACCG 638
Db 541 CTACAGAGCGCAAGAAATCAGGAGCTGCAGGCGCTTGCGGCGAGGCCATGCGGACCG 600
QY 639 GAATGAGTCCGGCTGAGAGCTGCTCAGAGCTTATTAACACAGCTGTCTCTTGAATGC 698
Db 601 GAATGAGTCCGGCTGAGAGCTGCTCAGAGCTTATTAACACAGCTGTCTCTTGAATGC 660
QY 699 GCGCTTCTCACCCCTGCCAGAGAGCTGCGGGCTCTTCTTCCACTGGTACGACTGCTTAC 758
Db 661 GCGCTTCTCACCCCTGCCAGAGAGCTGCGGGCTCTTCTTCCACTGGTACGACTGCTTAC 720
QY 759 TGGGGTCCCGGCGCCAGCAGCGTGCCTTGGCTGCGGAGAGGCGAGCTTCTTCAACAT 818
Db 721 TGGGGTCCCGGCGCCAGCAGCGTGCCTTGGCTGCGGAGAGGCGAGCTTCTTCAACAT 780
QY 819 CGGTGCGCTTCAACGCGAGATTTGGGGCGCGCCAGGACCGCTTCTGACCCGAGGGTGC 878
Db 781 CGGTGCGCTTCAACGCGAGATTTGGGGCGCGCCAGGACCGCTTCTGACCCGAGGGTGC 840
QY 879 CCGGCTATGAGAGCGCTTCCAGAGGCGGCTTGGGGCTTCCAGCCTCTCTGAGGGAGAACTT 938
Db 841 CCGGCTATGAGAGCGCTTCCAGAGGCGGCTTGGGGCTTCCAGCCTCTCTGAGGGAGAACTT 900
QY 939 CTCCCATGCGCGAGCGCAGACATGAGGGCTGCGCTTCCCTCTGCGCAGCTGAGGAGCTCAT 998
Db 901 CTCCCATGCGCGAGCGCAGACATGAGGGCTGCGCTTCCCTCTGCGCAGCTGAGGAGCTCAT 960
QY 999 GATGGCCCGAGCCAGGAGTGTGTTGAGGGCTCTTCAACACCTGCTCCATGGGCCCC 1058
Db 961 GATGGCCCGAGCCAGGAGTGTGTTGAGGGCTCTTCAACACCTGCTCCATGGGCCCC 1020
QY 1059 CCAAGACTGCTGCGCCAGCTGCGCTGCGGCGAGAGCGCCCGCAGGTTGGAGCGAGTA 1118
Db 1021 CCAAGACTGCTGCGCCAGCTGCGCTGCGGCGAGAGCGCCCGCAGGTTGGAGCGAGTA 1080
QY 1119 CAGGCTAGTGCACGGGACCATGGGCGCAGCGACCGGTCCACAGCTACGCTGTCTCTG 1178
Db 1081 CAGGCTAGTGCACGGGACCATGGGCGCAGCGACCGGTCCACAGCTACGCTGTCTCTG 1140
QY 1179 GACTGCGCTGTGTCATGTCAGAGCGGAGTACTTCCGCTCCCTGGGCCCACTACCGTAGC 1238
Db 1141 GACTGCGCTGTGTCATGTCAGAGCGGAGTACTTCCGCTCCCTGGGCCCACTACCGTAGC 1200
QY 1239 CATGGCCCTCTGCGAGCGGCTCCCGAGCGACCGAGGAGAGCTTCCCGACGCGAGCAGGT 1298
Db 1201 CATGGCCCTCTGCGAGCGGCTCCCGAGCGACCGAGGAGAGCTTCCCGACGCGAGCAGGT 1260


```
1299 CTTCTCTGCAGCCGCCACCTCTCTTAAGCCCGAGGCCCTGTGTCGCGCAGGAGCTGGA 1358
1261 CTTCTCTGCAGCCGCCACCTCTCTTAAGCCCGAGGCCCTGTGTCGCGCAGGAGCTGGA 1320
1359 GGAGCCGAGCAGCTTGGCAAGGCAACCTGAAGCGTGCATCTCTGGGCGAGGAGGCG 1418
1321 GGAGCCGAGCAGCTTGGCAAGGCAACCTGAAGCGTGCATCTCTGGGCGAGGAGGCG 1380
1419 GCTGCGGCTGTCAGCCCTGTGCGCGCTCTGTCGCGAGGTTGGACTGCTTGGGCTGTGAT 1478
1381 GCTGCGGCTGTCAGCCCTGTGCGCGCTCTGTCGCGAGGTTGGACTGCTTGGGCTGTGAT 1440
1479 CTCACAGCGCTGCAGCGCTCACTGCGCAAGTATGCGGAGCTCGACCGTGAAGTACATT 1538
1441 CTCACAGCGCTGCAGCGCTCACTGCGCAAGTATGCGGAGCTCGACCGTGAAGTACATT 1500
1539 CTGTGAGGCTGCGAGGCGCCGCGACATCCAGCGCTTAAGACCCACCAAGGAGGAGGAG 1598
1501 CTGTGAGGCTGCGAGGCGCCGCGACATCCAGCGCTTAAGACCCACCAAGGAGGAGGAG 1560
1599 GATGCCACGCTGTCAGCGGCGAGGCGCTGACATCTTCATTCGCTGGGCGCCCTGTC 1658
1561 GATGCCACGCTGTCAGCGGCGAGGCGCTGACATCTTCATTCGCTGGGCGCCCTGTC 1620
1659 TGTGTTCTCAGCAAGAACCGGTGCGGCTGCGGCGCTGCGGCGCCCTGACCTGACCGAGGAGA 1718
1621 TGTGTTCTCAGCAAGAACCGGTGCGGCTGCGGCGCTGCGGCGCCCTGACCTGACCGAGGAGA 1680
1719 GGGCGGCTTGGCTCAGCTTTCGCGGAGACTGCGCTGCTCTCATCGCTGCGCTCATTC 1778
1681 GGGCGGCTTGGCTCAGCTTTCGCGGAGACTGCGCTGCTCTCATCGCTGCGCTCATTC 1740
1779 AGGAGCCAGCGCGCGGCGCTGCGGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGG 1838
1741 AGGAGCCAGCGCGCGGCGCTGCGGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGG 1800
1839 GCAGCCATGAGGCTGCGAGACACCGGAGTGTGTCGAGGAGGAGGAGGAGGAGGAGGAG 1898
1801 GCAGCCATGAGGCTGCGAGACACCGGAGTGTGTCGAGGAGGAGGAGGAGGAGGAGGAGGAG 1860
1899 GGGCGGCGCCAGCTGCGAGGCTGCTGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGG 1958
1861 GGGCGGCGCCAGCTGCGAGGCTGCTGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGG 1920
1959 GGAGCGCGCGGCTGCTGCTGCGGCGCGCGGCGGCTTCTAAGGAGGAGGAGGAGGAGGAGGAG 2018
1921 GGAGCGCGCGGCTGCTGCTGCGGCGCGCGGCGGCTTCTAAGGAGGAGGAGGAGGAGGAGGAGGAG 1980
2019 TTGCAAGACCCCGGATCCAGTGGGCGAGTCCCGCGGCGGCTGCTCAACTGGAGCGGAGAA 2078
1981 TTGCAAGACCCCGGATCCAGTGGGCGAGTCCCGCGGCGGCTGCTCAACTGGAGCGGAGAA 2040
2079 GGCCAGCAGGCGAAGCTGAGGCTGCGGCGGCGGCTGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGG 2138
2041 GGCCAGCAGGCGAAGCTGAGGCTGCGGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGG 2100
2139 GCCTCATCTTGAAGCACCAGGCTGGCGG 2169
2101 GCCTCATCTTGAAGCACCAGGCTGGCGG 2131

RESULT 4
BC025767
LOCUS
DEFINITION Homo sapiens rhophilin, 3763 bp mRNA linear PRI 06-OCT-2003
clone MGC:34714 IMAGE:5211354, complete cds.
ACCESSION BC025767
VERSION BC025767
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

REFERENCE
AUTHORS

1 (bases 1 to 3763)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sánchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmitt, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257

JOURNAL
MEDLINE

2 (bases 1 to 3763)
Strausberg, R.
Direct Submission
Submitted (06-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

JOURNAL
PUBMED

12477932
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Hagnighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.D., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

REMARK
COMMENT

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 49 Row: j Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 19882236.

FEATURES
source

1. .3763
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:34714 IMAGE:5211354"
/tissue_type="Blood, adult leukocytes"
/clone_lib="NIH MGC_118"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1. .3763
/gene="RHPN1"
/note="synonyms: KIAA1929, RHPN, ODF5"
/db_xref="LocusID:114822"
102. .2114
/codon_start=1
/product="RHPN1 protein"


```

/misc_feature
/db_xref="AAH25767.1"
/db_xref="GI:19343951"
/db_xref="LocusID:114822"
/translation="MILERPDPGAGAEESPRLOQCDLSITQCGOLOSRRAIHHQI
DKELOMRTCAENLYRATSNRRVETVALEISVYNSNLOLLKEELELESGVDPGRHGS
EAVTVMPIGLKEYKELDSWFLKELISVHFGEDGASVEARIELEALROAMRTPSR
NESGJELLTYNQICFLDARFLTPARSLGLFPHWYDSLTGVPAQORALAPEKSGVLIF
NTGALHTQIGARQDSCTEGARAMEAFORAGAFSLLRENSHAFSPDMSAASLCAL
EQLMMAQACQCFEGLSPSPAPQDLAQLRAQEAAYAAEYVHRHTMAQPPVHD
YVPSWTALVHVKAERYLAHVMALCDGSPATEGELTHEQVFLQPTSSRPRG
PVLPOELRROLGKALHRAILGOEALRLHALCLRVLEVDLLRAVLSQTLQSLAK
YABLREDDFCAAEAPDIOPTKTHQPEARMPLRSQKGPDIPIHRLGPLSVFSKRNW
RVLGPHLLTRGEGGFLTGURDGPVLLIAAVIPCSQAAAGLKEGDIVSVNGOPCRWW
RAEYVTEUKAAGEAGASLOVSLPSLSDRRPVLGPRLLRSOREHGCKTP
ASTWASPRLLNWSRAQOGKTCGCPQPCAPVKPAPPSSLKHPGW"
204..341
/notes="H31; Region: H31 repeat"
/db_xref="CDD:pfam02185"
423..881
/notes="BRO1; Region: BRO1-like domain. This functionally
uncharacterised domain is found in a number of signal
transduction proteins, including Rhophilin and BRO1"
/db_xref="CDD:pfam03097"
1641..1841
/notes="PDZ; Region: Domain present in PSD-95, Dlg, and
ZO-1/2. Also called DHR (Dlg homologous region) or GLGF
(relatively well conserved tetrapeptide in these domains).
Some PDZs have been shown to bind C-terminal polypeptides"
/db_xref="CDD:smart00228"

```

ORIGIN

```

Query Match      89.9%; Score 1951; DB 9; Length 3763;
Best Local Similarity 100.0%; Pred. No. 7.5e-295;
Matches 1951; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 219 GGGGTGTACTCCGTGACGAGATCCAGTGGCGGACGTGCAGAGCCGCGAGGGCCAGAT 278
DB 161 GGGGTGTACTCCGTGACGAGATCCAGTGGCGGACGTGCAGAGCCGCGAGGGCCAGAT 220
QY 279 TCACAGCAGATTGACAGGAGCTGCAGATCGGACCGGGCGCTGAGAACCTTACAGAGC 338
DB 221 TCACAGCAGATTGACAGGAGCTGCAGATCGGACCGGGCGCTGAGAACCTTACAGAGC 280
QY 339 CACAGCAACAACCGGFTGAGAGAGACGGTGGCCCTGGAGCTGAGCTACGTCACACTCAA 398
DB 281 CACAGCAACAACCGGFTGAGAGAGACGGTGGCCCTGGAGCTGAGCTACGTCACACTCAA 340
QY 399 CTGCGAGCTGTGAGGAGGAGCTGAGAGAGCTGAGCGGTGGCGGTGGACCTTGGCGGCA 458
DB 341 CTGCGAGCTGTGAGGAGGAGCTGAGAGAGCTGAGCGGTGGCGGTGGACCTTGGCGGCA 400
QY 459 TGGAGCGAAGCTGTCACTGTCCCAATGATCCCTCGTGGCGCTGAGAGGACCAAGAGACT 518
DB 401 TGGAGCGAAGCTGTCACTGTCCCAATGATCCCTCGTGGCGCTGAGAGGACCAAGAGACT 460
QY 519 GGAATGCTTACCGCTGAGAGAGCTGATCTCAGTGCATCTTGGAGAGGAGCGGCGCTC 578
DB 461 GGAATGCTTACCGCTGAGAGAGCTGATCTCAGTGCATCTTGGAGAGGAGCGGCGCTC 520
QY 579 CTACAGAGCAGAAATCAGGAGCTGGAGCCCTCGGCGAGGCGCATGCGGACCCCGAGCG 638
DB 521 CTACAGAGCAGAAATCAGGAGCTGGAGCCCTCGGCGAGGCGCATGCGGACCCCGAGCG 580
QY 639 GAATGAGTCGGCGCTGAGAGCTGCTCAGAGCTATTACACAGAGCTGTGCTTCTGATGC 698
DB 581 GAATGAGTCGGCGCTGAGAGCTGCTCAGAGCTATTACACAGAGCTGTGCTTCTGATGC 640
QY 699 GGGCTTCTTCAACCCCTGCCAGAGCTCGGGCTCTTCTTCCACTGGTACAGCTCGCTTAC 758
DB 641 GGGCTTCTTCAACCCCTGCCAGAGCTCTGGGCTCTTCTTCCACTGGTACAGCTCGCTTAC 700
QY 759 TGGGGTCCCGGCCACGAGCGTGGCCCTTGGCGCTTTCGAGAGAGGCGAGCTTCTTCAACAT 818

```

```

DB 701 TGGGGTCCCGGCCACGACGCTGCTTGGCTTCGAGAGGGGAGGGTCTCTTCAACAT 760
QY 819 CGGTGCTCTCCACAGCAGATGTTGGGGCGCGCCAGAGACCGCTCTCTGACCGAGGGTGCCTCG 878
DB 761 CGGTGCTCTCCACAGCAGATGTTGGGGCGCGCCAGAGACCGCTCTCTGACCGAGGGTGCCTCG 820
QY 879 CGCGCTATGAGGGCTTCCAGAGGGCGCTCGGGGCTTTCAGCTCTCTGAGGGAGACTT 938
DB 821 CGCGCTATGAGGGCTTCCAGAGGGCGCTCGGGGCTTTCAGCTCTCTGAGGGAGACTT 880
QY 939 CTCCATGCGCGAGCCCGCAGATGAGCGCTCGCTTCCCTCTGCGCACTGGAGCAGCTCAT 998
DB 881 CTCCATGCGCGAGCCCGCAGATGAGCGCTCGCTTCCCTCTGCGCACTGGAGCAGCTCAT 940
QY 999 GATGGCCCGAGGCCAGGAATGTGTGTTGAGGGCTCTTCCACACCTGCTCTTCAATGCGCCC 1058
DB 941 GATGGCCCGAGGCCAGGAATGTGTGTTGAGGGCTCTTCCACACCTGCTCTTCAATGCGCCC 1000
QY 1059 CCAAGACTGCTGCTGGCCAGCTGCGCGCAGAGGGCGCGCCAGGTGCGACCGAGTA 1118
DB 1001 CCAAGACTGCTGCTGGCCAGCTGCGCGCAGAGGGCGCGCCAGGTGCGACCGAGTA 1060
QY 1119 CAGGCTAGTGACCGGACCATGCGCCAGCCAGCCCTCCACGACTAGCTGCTCTCTCTG 1178
DB 1061 CAGGCTAGTGACCGGACCATGCGCCAGCCAGCCCTCCACGACTAGCTGCTCTCTCTG 1120
QY 1179 GACTGCGCTGCTGATGTCAAGCCCGAGTACTTCCGCTCTCTGCGCCACCTACACGTA 1238
DB 1121 GACTGCGCTGCTGATGTCAAGCCCGAGTACTTCCGCTCTCTGCGCCACCTACACGTA 1180
QY 1239 CATGGCCCTCTGCGAGCGGCTCCCGAGCGACCGAGGGAGCTCCCGCCAGCCAGGAGGT 1298
DB 1181 CATGGCCCTCTGCGAGCGGCTCCCGAGCGACCGAGGGAGCTCCCGCCAGCCAGGAGGT 1240
QY 1299 CTCTCTGACGCGCCCGACCTCTCTAAGCCCGCGAGGGCGCTGTGCTGCGCGAGAGCTGGA 1358
DB 1241 CTCTCTGACGCGCCCGACCTCTCTAAGCCCGCGAGGGCGCTGTGCTGCGCGAGAGCTGGA 1300
QY 1359 GGAGCGAGGAGCTGCTGGCAAGGCACACCTGGAAGCTGCCATCTCTGGGGCAGAGGAGGC 1418
DB 1301 GGAGCGAGGAGCTGCTGGCAAGGCACACCTGGAAGCTGCCATCTCTGGGGCAGAGGAGGC 1360
QY 1419 GCTGCGGCTGACGCGCTGTGCGCGCTCTGCGCGAGGTGGAAGCTGCTGCGGGCTGTGAT 1478
DB 1361 GCTGCGGCTGACGCGCTGTGCGCGCTCTGCGCGAGGTGGAAGCTGCTGCGGGCTGTGAT 1420
QY 1479 CTCCAGAGCTGCGAGCGCTCAGCTGGCCAACTATGCGAGCTCGACCTGAGAGGAGTGA 1538
DB 1421 CTCCAGAGCTGCGAGCGCTCAGCTGGCCAACTATGCGAGCTCGACCTGAGAGGAGTGA 1480
QY 1539 CTGTGAGGCTGCGCGAGGCCCGGACATCCAGCTTAAGACCCACAGAAAGCCAGAGGCCAG 1598
DB 1481 CTGTGAGGCTGCGCGAGGCCCGGACATCCAGCTTAAGACCCACAGAAAGCCAGAGGCCAG 1540
QY 1599 GATGCGAGCTGCTCCAGGGGAGGGGCTGACATCTTCCATCGGCTGGGCGCCCTGTC 1658
DB 1541 GATGCGAGCTGCTCCAGGGGAGGGGCTGACATCTTCCATCGGCTGGGCGCCCTGTC 1600
QY 1659 TGTGTTCTCAGCAAGAACCGGTGGCGCTGGTGGGGCGCGCTTCCATCGCTGAGCGAGGA 1718
DB 1601 TGTGTTCTCAGCAAGAACCGGTGGCGCTGGTGGGGCGCGCTTCCATCGCTGAGCGAGGA 1660
QY 1719 GGGCGGCTTGTGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCC 1778
DB 1661 GGGCGGCTTGTGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCC 1720
QY 1779 AGGAGCGAGCGCGCGGCTGGCTGAGGAGGGCGGCTGAGGAGGGCGGCTGAGTGAATGG 1838
DB 1721 AGGAGCGAGCGCGCGGCTGGCTGAGGAGGGCGGCTGAGGAGGGCGGCTGAGTGAATGG 1780
QY 1839 GCAGCGCATGAGGTGGTGGAGACACGCGAGGTGGTGGAGAGGTGAGAGGTGCGGGAGA 1898
DB 1781 GCAGCGCATGAGGTGGTGGAGACACGCGAGGTGGTGGAGAGGTGAGAGGTGCGGGAGA 1840

```


1020 GACTGCCCTGTGTCATGTCAGAGCCGAGTACTTCCGCTCCCTGGCCCACTACACGCTAGC 1079
1239 CATGGCCCTCTGCGAAGGCTTCC-----1261
1080 CATGGCCCTCTGCGAGGCTCCGTCAGTGCCCAACACACTTGCCTCATGCTACTGCCAAG 1139
1262 -----CAGCGACCGAGGAGCTCCC 1283
1140 GCCGCCCGCGCAGGCTCACAGCTCTCTGTCCCCCAGCAGCCAGCGAGGAGCTCCC 1199
1284 CAGCGACGAGCGTCTTCTGAGCCGCCCACTCTCTTAAGCCCGCGAGCCCTGTGCT 1343
1200 CAGCGACGAGCGTCTTCTGAGCCGCCCACTCTCTTAAGCCCGCGAGCCCTGTGCT 1259
1344 GCCCGAGGAGCTGGAGGAGCGCAGGAGCTTGGCAAGGCAACCTGAAGCGTGCATCCT 1403
1260 GCCCGAGGAGCTGGAGGAGCGCAGGAGCTTGGCAAGGCAACCTGAAGCGTGCATCCT 1319
1404 GGGCGAGGAGCGCTGCGCTGACGCGCTGCGCGCTCTGCGCGCTGCGCGAGGTGACCT 1463
1320 GGGCGAGGAGCGCTGCGCTGACGCGCTGCGCGCTGCGCGCTGCGCGAGGTGACCT 1379
1464 GCTTCGGGCTGTGATCTCCAGACGCTGCGAGCGCTCACTGGCCCAAGTATCGGAGCTCGA 1523
1380 GCTTCGGGCTGTGATCTCCAGACGCTGCGAGCGCTCACTGGCCCAAGTATCGGAGCTCGA 1439
1524 CCGTGAGGATGACTTCTGTGAGGCTGCGAGCGCCCGGACATCCAGCTTAAGCCACCA 1583
1440 CCGTGAGGATGACTTCTGTGAGGCTGCGAGCGCCCGGACATCCAGCTTAAGCCACCA 1499
1584 GAAGCCAGGCGAGGATGACCGCTGCTCCAGGGAAGGGGCTGACATCTTCATCG 1643
1500 GAAGCCAGGCGAGGATGACCGCTGCTCCAGGGAAGGGGCTGACATCTTCATCG 1559
1644 GCTGGGCGCCCTGTGCTGTGTTCTCAGCAAGAACCGGTGGCGTGTGGGCGCGTCCA 1703
1560 GCTGGGCGCCCTGTGCTGTGTTCTCAGCAAGAACCGGTGGCGTGTGGGCGCGTCCA 1619
1704 CTTGACCGGAGGAGGCGGCTTGGCTCAGCTTCGGGAGACTGCGCTGCTCAT 1763
1620 CTTGACCGGAGGAGGCGGCTTGGCTCAGCTTCGGGAGACTGCGCTGCTCAT 1679
1764 CGCTGCGCTCATTTCCAGGAGCGCAGCGCGCGCTGCGCTGAGGAGGCGGCTCAT 1823
1680 CGCTGCGCTCATTTCCAGGAGCGCAGCGCGCGCTGCGCTGAGGAGGCGGCTCAT 1739
1824 TGTGTCACTGAATGGCGAGCCATGAGGTGTTGGAGACACCGGAGGTGTTGACGAGCT 1883
1740 TGTGTCACTGAATGGCGAGCCATGAGGTGTTGGAGACACCGGAGGTGTTGACGAGCT 1799
1884 GAAGGCTCGGAGAGGCGGCGCGCAGCTGAGGTGTTGGAGACACCGGAGGTGTTGACGAGCT 1943
1800 GAAGGCTCGGAGAGGCGGCGCGCAGCTGAGGTGTTGGAGACACCGGAGGTGTTGACGAGCT 1859
1944 ACTGCCAGCTTGGGGAGCCGCGCGCTGCTGCTGGGCGCCCGAGGGGCTTCTAAGGAG 2003
1860 ACTGCCAGCTTGGGGAGCCGCGCGCTGCTGCTGGGCGCCCGAGGGGCTTCTAAGGAG 1919
2004 CCAGAGGAGCATGTTGTCAGACCCCGGCATCCAGGTGGGCGAGTCCCGGCGCCCTCT 2063
1920 CCAGAGGAGCATGTTGTCAGACCCCGGCATCCAGGTGGGCGAGTCCCGGCGCCCTCT 1979
2064 CAACTGGAGCCGAAAGGCCAGGCGGAGACTGAGGCTGCCCGAGCTGTGGCCCTGCGCC 2123
1980 CAACTGGAGCCGAAAGGCCAGGCGGAGACTGAGGCTGCCCGAGCTGTGGCCCTGCGCC 2039
2124 AGTGAAGCCAGCTCGCCCTCATCTTTGAAGCAACCCAGGGTGGCG 2169
2040 AGTGAAGCCAGCTCGCCCTCATCTTTGAAGCAACCCAGGGTGGCG 2085

AB067516 3647 bp mRNA linear PRI 15-SEP-2001
Homo sapiens mRNA for KIAA1929 protein, partial cds.
AB067516
AB067516.1 GI:15620916
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Nagase, T., Kikuno, R. and Ohara, O.
Prediction of the coding sequences of unidentified human genes.
XXI. The complete sequences of 60 new cDNA clones from brain which
code for large proteins
DNA Res. 8 (4), 179-187 (2001)
21456161
11572484
2 (bases 1 to 3647)
Chara, O., Nagase, T. and Kikuno, R.
Direct Submission
Submitted (27-JUL-2001) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
232-0812, Japan (E-mail:cdnainfo@kazusa.or.jp.
URL: http://www.kazusa.or.jp/huge, Tel: 81-438-52-3913,
Fax: 81-438-52-3914)
Location/Qualifiers
1. 3647
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="fj08985"
/tissue_type="brain"
/note="vector:pBluescriptII SK plus"
1. 3647
/gene="KIAA1929"
<1. 1233
/gene="KIAA1929"
/note="Start codon is not identified. fj08985 cDNA clone
for KIAA1929 has 349-bp, 320-bp and 413-bp insertions
after the positions 450, 936 and 1283 of the sequence of
KIAA1929, respectively."
/codon_start=1
/product="KIAA1929 protein"
/protein_id="BAB67822.1"
/db_xref="GI:15620917"
/translation="EERPDAAGAEESPRLOGCDLSLTQIQCGQLQSRRAQIHOQIDKE
LQMTGAENLYRATSNRVRVETVALSLSYVNSLQQLKELELGGVDPGRHSEAV
TVMPLGLKETKELDWSTPLKELISVHFEGDGA YEARELEALFOAMTPGRNES
GLELLTAYTQLCFLDARFLTPARSLGLFHHYDLSLTCVPAQQRALAFKGSVLFNIG
ALHTQIGARDRSCTEGARRAEAFQRAAGAFSLLENFSAFSPDMSAALCALEQL
MMAQACVCEGLSPSPAPQDCLQLRLAQEAQVAAYRLVHRTVAQPPVHDYVP
VSWTALVHVKAEPFRSLAHVHVAVALCDGSRCPPLPMLVLPFRPRAGLTASLSPSSD
RGRAPHARAGLPAAPHLL"
ORIGIN
Query Match 86.0%; Score 1865; DB 9; Length 3647;
Best Local Similarity 96.3%; Pred. No. 2e-281;
Matches 1951; Conservative 0; Mismatches 0; Indels 76; Gaps 1;
QY 219 GGGCTGTGCTCCCTGACGCGAGATCCAGTGGCGGCGAGCTGCAGAGCCGCGAGGCCCGCAGAT 278
DB 51 GGGCTGTGCTCCCTGACGCGAGATCCAGTGGCGGCGAGCTGCAGAGCCGCGAGGCCCGCAGAT 110
QY 279 TCACCAGCAGATTGACAGGAGCTGCAGATCGGACGGCGCTGAGAACCTCTACAGAGC 338
DB 111 TCACCAGCAGATTGACAGGAGCTGCAGATCGGACGGCGCTGAGAACCTCTACAGAGC 170
QY 339 CACCAGGACAAACCGGGGTGAGAGACCGGTGCGCCCTGGAGCTGAGCTCACTCAAC 398
DB 171 CACCAGGACAAACCGGGGTGAGAGACCGGTGCGCCCTGGAGCTGAGCTCACTCAAC 230
QY 399 CCTGCAGCTGCTGAAGGAGGAGCTGCAGGAGCTCAGCGGTGCGGTGGACCCCTGGCGGCA 458

of Fac. Med., Yoshida Konoe-Cho Sakyo-Ku, Kyoto City, 606, Japan	
Location/Qualifiers	
1. .1932	
/organism="Mus musculus"	
/mol_type="mRNA"	
/db_xref="taxon:10090"	
1. .1932	
/function="small GTPase Rho binding protein"	
/codon_start=1	
/product="rhophilin"	
/protein_id="AAC523388.1"	
/db_xref="GI:1176422"	
/translation="MIIEKRPDQGTGESSRRQDDGSIKRGYSFVONOPGLOSHR ARLHQISXELMRITGAENRATSNWRETVALELSYVNSLQILLKEELALSTSV DVDPQEGGTTIMIPLIGKETKELDWATPLKELISEHFGEDGTSFETIEQLELDLRQ ATPTPSRDAAGLLAAYSQLDFARFSPSRFGLLPHFYDLSLTGVPAAQALAF EKGSLVFNIGALHTOIGARODSCSTETGNAHAEAFORAFARLLRENFSHAPSOMS AASLMEQLMIAQOEICIFKGLLPASATPDICPOLQLOAAAVATRGVLFVHRAM AQPPVRDYLPSATNINLHVXAEHFCALAHYHAAALCESHPAKGELAROHVFOPTP HEPLGPTLPQHPDRRLAKAHLKRLGGQEALRLHTLCRLVKRLDQLGVVYQALR RSLAKYSQLEEDDFEATEAPDIQPKHTQTPGPIUSVFTKMRQLVGPVHVRGEG GFGLTLRGDSPLVIAAVPGGQAESGLKEGDYIVSYNGQPCKWMKHELVYTVQRSMG EEGVSLQVSLSPPEPRGTGPRRAALLMNORECGFETPMPTTRTPWPIILGWSRKNQ GKTGSHDPDCTNRCVTCp"	
ORIGIN	
Query Match	50.8%; Score 1102.2; DB 10; Length 1932;
Best Local Similarity	75.9%; Pred. No. 2,2e-162;
Matches 1450; Conservative	0; Mismatches 383; Indels 78; Gaps 4;
Qy	219 GGGCTGTGATCCCTGACGCGNATCCAGTGGCGGACGTGCAGAGCGCGAGGGCCAGAT 278
Db	81 GGGCTATGGTCCCTTTGTGAAAACACGACCTGGCCAGTGCAGAGTCAAGGGTCCGGT 140
Qy	279 TCACACAGAGATTGACAGGAGCTGCAGATCGGACCGGGCGCTCAGAAACCTCTACAGAGC 338
Db	141 CCACACAGATCAGCAAGGAGCTCCGGATCGGACCGGTCGTGAAACCTCTACAGAGC 200
Qy	339 CACACAGAAACCCGGGTGAGAGAGACGGTCCGCTCGAGCTGAGCTACGTCAACTCCAA 398
Db	201 CACCAGCAACACCTGGGTCCGAGAAACAGTGCACACTCGAGCTGAGCTACGTCAACTCAA 260
Qy	399 CCTGCAGCTGTGAAGGAGGAGCTGGAGGAGCTCAGGGTGGCTGGACCTCGCCCGGCA 458
Db	261 CCTGCAGCTGTGAAGGAGGAGCTGGAGGAGCTCAGGACCTGAGGAGTGTGAGACAGCC 320
Qy	459 TGGGAGGCAAGCTGTCACTGTCCCATGATCCCTCGGGCTGAAGGAGACCAAGGAGCT 518
Db	321 AGAGGCGAAGGGATTACTATTCCATGATCCCCCTGGGGCTGAAGGAAACCAAGGAGTT 380
Qy	519 GGAATGCTTACACCGCTGAAGGAGCTGATCTCAGTGCACTTTGGAGAGGACGGGCGCTC 578
Db	381 GGAATGCGCCACACCGCTGAAGGAGCTGATCTCAGCACCTTTGGAGAGGATGTTACCTC 440
Qy	579 CTACAGGACAGAAATCAGGAGGAGCTGGAGGCGCTGGGCGAGCGCATCGGACCCCCAGCGG 638
Db	441 CTTTGAACAGAGATTCAGGAACTAAGAAGACTACAGGAGGCGCCAGAGACTCCAGCAG 500
Qy	639 GAATGAGTCCGGGCTCGAGCTGCTCAGACCTATTACAAACAGAGCTGTGCTTCCTGGATGC 698
Db	501 GGACGAGGCGGCGCTGGATCTGCTCGACGCTACTATAGCCAGCTCTGTTCCTGGAGCGC 560
Qy	699 CGCTTCTCACCCTGCCAGGAGCTCGGGCTCTTCTCCACTGGTACGACTCGCTTAC 758
Db	561 ACCTTCTCTAGTCCCTCCAGAGCGCCGGGCTGCTCTTCACCTGGTACGACTCGCTCAC 620
Qy	759 TGGGTCCCGCCACGACGCGTGCCCTCGGCTTCAGAGAGGCGAGCGTTCTCTTCAACAT 818
Db	621 GGGGTCCCGCCACGACGAGCGGCTCTGGCTTCAGAGAGGCGAGTGTACTGTTAACAT 680
Qy	819 CGGTGCTCCACACGACGAGATTGGGGCGCGCAGAGACCGCTCTCTGCAACGAGGTGCCG 878
Db	681 TGGGCACTCCACACGACGATTGGAGCAGCGGAGGATTGCTCCTCAGCCGAGGGACCA 740

Qy	879 CCGCGCTATGAGGCGCTTCCAGAGGCGCGCTGGGGCGCTTTCAGCGCTCTCTGAGGAGAACTT 938
Db	741 CCACGCTGCTGAGGCTTTCCAGAGGCGCTGCTGGGGCGCTTTCAGACTCTCTGAGGAGAACTT 800
Qy	939 CTCCTCATGGCGCGAGCCGAGACATGAGCGCTGCTCCCTCTCTGCGCATGAGGAGCTCAT 998
Db	801 CTCCTCATGGCGCGAGCCGAGACATGAGTGTGCTCTCTCTCTCCATGCTGAGGAGCTCAT 860
Qy	999 GATGGCCGAGCCCGAGAAATGTGTGAGGGCTCTTACACACCTGCTCCATGCGCC 1058
Db	861 GATTGGCCGAGCCGAGAAATGTCTTCAAGGGCGCTTCTACTGCCAGCTCTTCCCAACC 920
Qy	1059 CCAAGAGCTGCTGCGCCAGCTGCGCTGGCGCAGAGGAGCGCCCGAGGTGCGAGCCGAGTA 1118
Db	921 TGACATCTGTCCAGATCAACTGAGCTGCGCCAGGAGCTGCCAGGTGCAACTGAATA 980
Qy	1119 CAGGCTAGTGACACCGAGACCATGCGCCAGCCGCTCCAGCATAGCTGCTGTCTCTG 1178
Db	981 TGGGCTTGTGACCGGGCCATGGCCAGCCACTGTCCGAGACTTACCTGCTGCTCTCTG 1040
Qy	1179 GACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1238
Db	1041 GACTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1100
Qy	1239 CATGGCCCTCTGCGACCGGCTCCCAGCGACCGAGGAGAGCTCCCCACGCAACGAGCAGGT 1298
Db	1101 CATGGCCCTCTGCGAAGCCACCCAG---CCAGGGTGAACCTCGAAGGCGAGGACAGCT 1157
Qy	1299 CTTCTGTCAGCCCGCCCTCTCTTAAGCCCGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTG 1358
Db	1158 CTTCT---CAGCCCTTCCACTCCCTGAGCCCTGGGCGCCACACTGCGCGAGCACCCAGA 1214
Qy	1359 GAGCGCAGGAGCTTGGCAAGGACACACTGAGCGTGCATCTCTGGGGCAGGAGGAGGC 1418
Db	1215 AGACCGCAGGAGGCTTGGCAAGGCGCCACCTGAGCGTGCATCTCTGGGGCAGGAGGAGGC 1274
Qy	1419 GCTGCGGCTGCAAGCCCTGTGCGCGCTCTGCGCGAGTGCAGCTGCTGCTGCTGCTGCTG 1478
Db	1275 GCTGCGGCTTCAACACTTGTGCGAGTCTCTGCGCAAGTGTGACCTGCTGCAAGTGTGGT 1334
Qy	1479 CTCCTCAGAGCTGAGCGCTCACTGGCGCAAGTATGCGGAGCTCCACCTGAGGATGACTT 1538
Db	1335 GACTCAGGCACTGCGGCGCTCTCTGCGCAAGTACTCAGCTCAGCGCGAGGAGTACTT 1394
Qy	1539 CTGTGAGGCTGCCGAGGCGCCGAGCATCCAGCGCTTAAGACCCACAGAGCCAGAGGCGCAG 1598
Db	1395 CTTGAGGCTTACGAGGCGCCCGAGCATCCAAACCCAGACTCACCAGACCCCGA----- 1448
Qy	1599 GATGCCACGCTGTCTCCAGGGGAAAGGGGCTTGACATCTTCCATCGGCTGGGGCGCTGTC 1658
Db	1449 -----GGGACCCCTTGTCT 1460
Qy	1659 TGTGTTCTCAGCAAGAAACCGTGTGGCGCTGCTGGGGCGCTCCACCTGACCCGAGGAGA 1718
Db	1461 TGTGTTTTCACCAAGAAACCGATGCGCAGTGTGTGAGGGCGCGTGCACTGATCTCAGGAGA 1520
Qy	1719 GGGGCGCTTTTGGCTCAGCTTTCGGGGAGACTCGGCTGCTCTCTCTCTCTCTCTCTCTCT 1778
Db	1521 GGGAGGCTTTCGGGTTTCACTGCTGCGGGAGACTCTCTCACTCTCTCTCTCTCTCTCTCT 1580
Qy	1779 AGGAGCCAGCCCGCGCGCTGCGCTGAAAGAGGCGGCTACATTTGTTGCTGTAATGG 1838
Db	1581 TGGGGGTGAGGCTGAGTCCGCTGCGCTGAGGAGGCGGCTTACATCTGCTGCTGCTGCTGCT 1640
Qy	1839 GCAGCATGAGGCTGAGTGGAGACCGCGAGGCTGTCAGGAGCTGAGAGGCTGCGGGGAGA 1898
Db	1641 GCAGCGTGCAGTGGTGGAGAGCCTGAGGCTGAGAGCAGCTGAGGAGCATGGGCGA 1700
Qy	1899 GCGGGCGCCACGCTGAGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1958
Db	1701 GGAAGGCTGAGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1760

QY 1959 GGACCGCGCGCCCTCTCTGCTGGCGCCCGACGGGGCTTCTAAGGAGCCAGAGGAGCATGG 2018
 Db 1761 ACCCGGCGGACACATGCTG-----TGGAACCCAGAGAGTGTGG 1802
 QY 2019 TTGCAAGACCCCGGCATCCACGTGGCGGAGTCCCGGCCCTCTCAACTCGAGCCGAAA 2078
 Db 1803 TTTTGAGACACCGATGCCACACGACGACCGCGCCCTGCGCCATCTTGGCTGGAGCCGCAA 1862
 QY 2079 GGCCCGACAGGCGAAGCTGAGAGCTGCGCCCGACCCCTGTGCCCCAGTGAA 2129
 Db 1863 GAATAAACAAGGCAAGACTGGAAGCCACCTGACCCCTGCACCAACAGGAA 1913

RESULT 8
 BC052010
 LOCUS
 DEFINITION Mus musculus rhophilin, Rho GTPase binding protein 1, mRNA (CDNA
 Clone MGC:62337 IMAGE:5717786), complete cds.
 BC052010
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2301)
 Strausberg R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zerbey, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, D.,
 Stappenko, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Frange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahy, J., Helton, E., Ketterman, M., Madan, A., Rodriguez, S.,
 Sanchez, A.A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Stalska, U., Smalls, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 12477932
 2 (bases 1 to 2301)
 Strausberg, R.
 Direct Submission
 Submitted (01-MAY-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 cDNA Library Preparation: Dr. Jim Lin, University of Iowa
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
 Series: IRAK Plate: 114 Row: k Column: 5
 This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 6680084.
 Location/Qualifiers
 1..2301
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="MGC:62337 IMAGE:5717786"
 /tissue_type="Brain, mouse 12.5 dpc"
 /clone_lib="NIH_BMAP_FDO"
 /lab_host="DH10B"
 /note="Vector: pYX-ASC"
 1..2301
 /gene="Rhpn1"
 /note="synonym: Rhophilin"
 /db_xref="LocusID:1478"
 /db_xref="MGI:1098783"
 66..1421
 /codon_start=1
 /product="Rhpn1 protein"
 /protein_id="AAH52010.1"
 /db_xref="GI:30354357"
 /db_xref="LocusID:1478"
 /translation="MILERPDDGGTGESSRRPDDGSIRKYGVSFVONOPGQLOSHR
 ARLHQQISKELRMRTGAENLYRATSNWTRETVALLSLVYNSNLIQLKEELAEISTSV
 DVDQPGEGITPMIPLGLKETKELDMATPLKELISEHFGEDGTSFETHELELDLRQ
 ATRPSRDRAGLDLAAYYSQCLFLDARFSPRSFGLLFWYDTSUTGVPAQORALAF
 EKVSVLFNGLHTQAGQDCSTEGTNHAAFAFORAAGAFELLENESHAFSPDMS
 AASLMEQLMIAQOECIFKGLLPASATPDIQDQLAQEAQAVATYGLVHRAM
 APPVRYLPASNTNLRAVKAHFALAHYAAMALCESHPAKGELARQEHVFPSTP
 HPTPLTPHPEDRRKLGGAAASHLVQSPAQGGPAAGCGDSGTAALPQVLTARAR
 G"
 189..404
 /note="HR1; Region: HR1 repeat"
 /db_xref="CDD:pfam02185"
 408..866
 /note="BR01; Region: BR01-like domain. This functionally
 uncharacterised domain is found in a number of signal
 transduction proteins, including Rhophilin and BR01"
 /db_xref="CDD:pfam03097"
 misc_feature 189..404
 /note="HR1; Region: HR1 repeat"
 /db_xref="CDD:pfam02185"
 408..866
 misc_feature
 /note="BR01; Region: BR01-like domain. This functionally
 uncharacterised domain is found in a number of signal
 transduction proteins, including Rhophilin and BR01"
 /db_xref="CDD:pfam03097"
 ORIGIN
 Query Match 47.8%; Score 1036.6; DB 10; Length 2301;
 Best Local Similarity 73.3%; Pred. No. 3 6e-152;
 Matches 1463; Conservative 0; Mismatches 389; Indels 143; Gaps 5;
 QY 219 GGGCTGTGACTCCCTGACGACGATCCAGTGGCGGCGAGCTGCAGAGCCGAGGGCCGACAT 278
 Db 146 GGGCTATGGCTCTCTTGTGCAAAACAGCCTGGCCAGCTGCAGAGTCAACAGGCTCGGCT 205
 QY 279 TCACCAGCAGATTGACAAAGGAGCTGCAGATGGGAGGGCGGTGAGAACTCTACAGAGC 338
 Db 206 CCACCAGCAGATTCAGCAGGAGCTCCGATGGGACCGGTGCTGAAAACCTCTACAGAGC 265
 QY 339 CACCAGCAACACCGGGTGAGAGACGCTGCCCTGGAGCTGAGCTACGCTCAACTCCAA 398
 Db 266 CACCAGCAACCTGGTCCGAGAACAGTGCATCGAGAGCTAGCTAGTCACTCCAA 325
 QY 339 CTGCACTGCTGAAGGAGGAGCTGAGAGGAGCTCAGCGGTGGCGGCTTGGCGCGCA 458
 Db 326 CTGCACTGCTGAAGGAGGAGCTGAGAGGAGCTGAGAGGAGCTGAGAGGAGGAGT 385
 QY 459 TGGGAGCGAAGCTGTCACTGCTCCCATGATCCCTGGGCTGAGGAGGAGGAGGAGCT 518
 Db 386 AGAGGCGAAGGAGTACTATTCCATGATCCCCCTGGGCTGAGGAGGAGGAGGAGT 445
 QY 519 GGACTGTGCTACACCGCTGAAGGAGCTGATCTCAGTGCACCTTTGAGAGGAGGAGGAGCT 578
 Db 446 GGACTGGGCCACACCGCTGAAGGAGCTGATCTGAGCAGCTTTGGAGGAGGAGTACCTC 505
 QY 579 CTACGAGCGAAGATCAGGAGGAGCTGGAGGCGCTTCGGGAGGCCCATGGGACCCCGAGCG 638
 Db 506 CTTTGAGACAGAGATCCAGGAACCTAGAAGACCTACCGGACGCCACGAGGAGCTCCAGCAG 565

QY	639	GAATGAGTGGGCTGGAGCTGTCAAGCCTATTACACAGCTGTGCTTCTGATGC	698
Db	566	GGACGAGGGGGCTGGATCTGCTCGCAGCTACTATAGCCAGCTCTGTTCCTGACGC	625
QY	699	GCCTTTCTTCACTTCGACAGGAGCTTCGGCTTCTTCCACTGTGTACACTCGCTTAC	758
Db	626	ACGTTTCTTCACTTCGACAGGAGCTTCGGCTTCTTCCACTGTGTACACTCGCTTAC	695
QY	759	TGGGCTTCCGCGCCAGCAGCTGTGCTTGGCTTTCGAGAGGGAGCGCTTCTTCAACAT	818
Db	686	GGGGTCCCAGCAGCAGCGGGCTCTGGCTTTCGAGAGGGAGCTTACTTAACT	745
QY	819	CGTGCCCTTCCACACGAGATTGGGCGCGCAGGACCGTCTCTGACACGAGGTGCGG	878
Db	746	TGGGGCACTCCACACGAGATTGGAGCACGCGAGATTGCTCTGCAACCGAGGACAA	805
QY	879	CCGGCTATGAGGCTTCCAGAGGGCGCTGGGCTTTCAGCTCTCTGAGGGAGAACTT	938
Db	806	CCAGCTGTGAGCTTTCAGAGGGCTGTGGGCTTTCAGACTCTCTGAGGAGAACTT	865
QY	939	CTCCATGCGCGAGCCAGACATGAGCGCTGGCTTCTCTGCGCACTGGAGAGCTCAT	998
Db	866	CTCCACGCGCAGACCCAGACATGAGTGTGCTCTCTCTCCATGTGGAGCAGCTCAT	925
QY	999	GATGGCCAGCCAGGAATGTGTGTTGAGGCTCTCACCACTGCTCCATGGCCCC	1058
Db	926	GATTGCCAGGCA CAGGAATGCATTTCAAGGGCTTTTACTGCCAGCTCTGCCACACC	985
QY	1059	CCAGACTGCTGCGCCAGCTGCGCTGGCGCAGAGGCGCCCGAGGTGGCAGCGAGTA	1118
Db	986	TGACATCTGTCCAGTCAACTGCACTGTGCCCCAGAGAGCTGCCAGGTGGCACTGAATA	1045
QY	1119	CAGGCTAGTGACCGGACCATGGGCCAGCCACCCGTCACACTACAGTGTCTCTCTG	1178
Db	1046	TGGGCTTGTGACCGGGCCATGGCCAGCCACCTGTCCGAGACTACCTGCTGCTCTCTG	1105
QY	1179	GACTGCTGTGTGATGCAAGGCGAGTACTTCGGCTCCTGGCCACTACCACTGAGC	1238
Db	1106	GACTAACTTGGCGCAGCTCAAGGTGAGCACTTCTGGCCCTTGGCCCACTATCATGACG	1165
QY	1239	CATGGCTCTGGCAGGCTCCCGACGACCGAGGAGCTCCCGCAGCAGCAGAGGT	1298
Db	1166	CATGGCTCTGGCAGGACCCAG---CCAGGGTGAATCTCGAAGCGAGGACAGT	1222
QY	1299	CTTCTGACGCCCCCACTCTTAAGCCCGAGGCGCTGTGCTGCGCGAGGAGCTGGA	1358
Db	1223	CTTC---CAGCCCTCCACTCCCATGAGCCCTGGGCGCCACACTGCGCGAGCACCAGA	1279
QY	1359	GGAGCGAGGAGCTTGGCAGGACACCTGAGCGTGCCATCTCTGGGCGAGGAGGCG	1418
Db	1280	AGACCGAGGAGCTT-----GGAGGAGCG	1304
QY	1419	GCTGGGCTGACGCGCTGTGCGGCTGCTGCGCAGGTGGAACCTGCTTGGGCTGTGAT	1478
Db	1305	GCTGGGCTTACACTTGTGCAAGTCTCTGGCAAGGTGGACCTGCTGCAGTGTGTGT	1364
QY	1479	CTCCAGACGCTGACGCGCTCACTCGGCAAGTATGCGGAGCTGCAACGTTGAGGATGACTT	1538
Db	1365	GACTCAGGCACTGCGGCGCTCTGCGCAAGTACTCACAGCTGCGCGCGAGGATGACTT	1424
QY	1539	CTGTGAGGCTGCGAGGCGCGGACATCCAGCTTAAGACCCACAGAGCCAGAGGCGCAG	1598
Db	1425	CTTCAGGCTACGAGGCGCGGACATCCACCCAGACTCACAGACCCCGAGGAGTGA	1484
QY	1599	GATGCCAGCTGTCTCCAGGGAAGGGCTTGACATCTTCCATGCGCTGGGCGCTCTGTC	1658
Db	1485	GATGCTAGCTGTGCGAGGTGAGGTGACTGACTTCTTCCATCGCTGGGACCTTGTGTC	1544
QY	1659	TGTGTTCTCAGCCAGAACCGGTGGGCTGTGGGCGCGCTCCACTTACCCGAGGAGA	1718
Db	1545	TGTGTTTTTCAACCAAGAACCGAGTGTGTAGGGCGGTGTCATGACTCGAGGAGA	1604

QY	1719	GGCGGCTTTGGCTTCAAGCTTTCGGGAGACTCGCTGTGCTTCACTCGTCCCTCATTC	1778
Db	1605	GGGAAGCTTCGGGTTTCAAGCTTTCGGGAGACTCTCCAGTCTTCACTCGTCTGCTGCTCC	1664
QY	1779	AGGAGCCAGGCGCGCGG-----	1795
Db	1665	TGGGGTTCAGGCTGAGGTAAAGCCCTCTAGCCCCGGGTGCTGGAATCTTAGAGGACCGC	1724
QY	1796	-----CGGCTGGCTTGAAGGAGGG	1814
Db	1725	TAGATATGCCACCACTTCCAGATTCATTTCTTCGCACTCGCTGGCTTGAAGGAGGG	1784
QY	1815	CGACTACTTGTGTGAGTGAATCGGAGCAGCATGAGGTGTGTGAGACACCGGAGAGTGT	1874
Db	1785	CGACTACTTGTGTGAGTGAATCGGAGCAGCATGAGGTGTGTGAGACACCGTGGAGGTGT	1844
QY	1875	GAGGAGCTTGAAGCTTGGGAGAGGCGGCGCAGCTGCAAGTGTGTGCTGTGCTGCC	1934
Db	1845	GACACAGCTGAGGAGCATGGGCGAGGAGCGTGTGAGCTGTGAGTGTGTGCTGTGCTGCC	1904
QY	1935	CAGCTCTAGACTTGGGAGACCGCCCGCTGCTGCTGCTGGGCCCCAGGGGCT	1994
Db	1905	CAGCTCTAGACTTGGGAGACCGCCCGCTGCTGCTGCTGGGCCCCAGGGGCT	1950
QY	1995	TCTAAGAGCAGGAGGAGCATGTTTGAAGACCGCGGATCCACGTGGGCGAGTCCCGG	2054
Db	1951	-----TGAACCGAGAGAGTGTGTTTGAAGACCGCGGATCCACGTGGGCGAGTCCCGG	2006
QY	2055	GCCCTCTCTCACTGGAGCGGAGGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG	2114
Db	2007	GCCCTCTCTCTGGTGGAGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG	2066
QY	2115	CTGTGCCCCAGTGAA	2129
Db	2067	CTGACCAACAGGAA	2081

RESULT	9
BD183400	
LOCUS	BD183400
DEFINITION	Novel genes and proteins encoded by the genes.
ACCESSION	BD183400.1
VERSION	GI:31875600
KEYWORDS	JP 2002345492-A/113.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	1 (bases 1 to 4729)
AUTHORS	Ohara, O., Nagase, T. and Nakajima, D.
TITLE	Novel genes and proteins encoded by the genes
JOURNAL	Patent: JP 2002345492-A 113 03-DEC-2002;
COMMENT	KAZUSA DNA RESEARCH INSTITUTE
	OS Homo sapiens (human)
	PN JP 2002345492-A/113
	PD 03-DEC-2002
	PF 26-FEB-2002
	PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
	PC C12N15/09, C07K14/47//A61K31/711, A61K38/00, A61K48/00, A61P25/00,
	PC A61P25/14, A61P35/00, C12N15/00, A61K37/02
	PC A61P25/18, A61P35/00, C12N15/00, A61K37/02
	CC Novel genes and proteins encoded by the genes FH Key
	Location/Qualifiers
	FT CDS (518).. (1540).
FEATURES	Location/Qualifiers
source	1..4729
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
ORIGIN	
Query Match	36.8%; Score 798.8; DB 6; Length 4729;
Best Local Similarity	92.1%; Pred. No. 3.7e-115;

Matches		853;	Conservative	0;	Mismatches	72;	Indels	1;	Gaps	1;
Qy	1244	CCCTCTGCGACGGCTCCCGAGGACCGAGGAGAGTCCCCACGACGACGAGCTCTTCC	1303							
Db	2235	CCCGTGGCGCGACCCACCAAGAAAGTGGCTGTGATGAGCCCAACAGCCCTGGCGTGC	2294							
Qy	1304	TGCAGCCCCCACCCTCTTAAGCCCGAGGCCCTGTGCTGCGCAGGAGCTGGAGGC	1363							
Db	2295	CGACTCTTCTGCGAGCTCCAGGGCCCAAGGGCCACATGGTGTGTGACATCCCATGC	2354							
Qy	1364	GCAGGAGCTTGGCAAGGCACACCTGAAGCGTGCCTCTCTGGGGCAGGAGGCGCTGC	1423							
Db	2355	CCGCGTGC-AGGCAAGGCACACCTGAAGCGTGCCTCTCTGGGGCAGGAGGCGCTGC	2413							
Qy	1424	GGCTGCAGCCCTGTGCGGCTCTGCGGAGGTGAGCTGCTTCGGGCTGTGATCTCC	1483							
Db	2414	GGCTGCAGCCCTGTGCGGCTCTGCGGAGGTGAGCTGCTTCGGGCTGTGATCTCC	2473							
Qy	1484	AGACGCTGCAGCGCTCACTGGCCCAAGTATGCGGAGCTCGACCGTGAAGTACTTGTG	1543							
Db	2474	AGACGCTGCAGCGCTCACTGGCCCAAGTATGCGGAGCTCGACCGTGAAGTACTTGTG	2533							
Qy	1544	AGGCTGCCAGGCCCGGACATCCAGCCTAAGACCCACAGAGCCAGAGCCAGCATGC	1603							
Db	2534	AGGCTGCCAGGCCCGGACATCCAGCCTAAGACCCACAGAGCCAGAGCCAGCATGC	2593							
Qy	1604	CAGGCTGTCCAGGGGAAGGGCTTGACATCTTCCATCGGCTGGGCCCCCTGTCTGTG	1663							
Db	2594	CAGGCTGTCCAGGGGAAGGGCTTGACATCTTCCATCGGCTGGGCCCCCTGTCTGTG	2653							
Qy	1664	TCTCAGCAAGAACCGTGGCGGCTGGTGGGCGCGTCCACCTGACCCAGAGAGGGCG	1723							
Db	2654	TCTCAGCAAGAACCGTGGCGGCTGGTGGGCGCGTCCACCTGACCCAGAGAGGGCG	2713							
Qy	1724	GCTTTGGCTCACGCTTCGGGAGACTCGCCTGTCTCTATCGCTGCGCTTCACAGGA	1783							
Db	2714	GCTTTGGCTCACGCTTCGGGAGACTCGCCTGTCTCTATCGCTGCGCTTCACAGGA	2773							
Qy	1784	GCCAGCGCGCGCGCTGGCTGAAGAGGGGCGACTACATTTGTCAGTGAATGGCGAGC	1843							
Db	2774	GCCAGCGCGCGCGCTGGCTGAAGAGGGGCGACTACATTTGTCAGTGAATGGCGAGC	2833							
Qy	1844	CATGACGCTGGTGGAGACACGCGGAGGTGGTCAAGGAGTGAAGCTGCGGAGAGGGC	1903							
Db	2834	CATGACGCTGGTGGAGACACGCGGAGGTGGTCAAGGAGTGAAGCTGCGGAGAGGGC	2893							
Qy	1904	GCGCCAGCTCGAGTGGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACC	1963							
Db	2894	GCGCCAGCTCGAGTGGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACC	2953							
Qy	1964	GCGGCGGCTCTGCTGGGCGCCAGGGGCTTTAAGGAGCGAGGAGCATGTTGCA	2023							
Db	2954	GCGGCGGCTCTGCTGGGCGCCAGGGGCTTTAAGGAGCGAGGAGCATGTTGCA	3013							
Qy	2024	AGACCCCGGATCAACGCTGGGCGAGTCCCGGCGCTCTCACTGAGCGCGGAGGCGC	2083							
Db	3014	AGACCCCGGATCAACGCTGGGCGAGTCCCGGCGCTCTCACTGAGCGCGGAGGCGC	3073							
Qy	2084	AGCAGGGCAAGCTGGAGGCTGCCCCAGCCCTGTGCCCCCAGCTGAAGCAGCTCGGCCCT	2143							
Db	3074	AGCAGGGCAAGCTGGAGGCTGCCCCAGCCCTGTGCCCCCAGCTGAAGCAGCTCGGCCCT	3133							
Qy	2144	CATCTTTGAAGCACCCAGGGTGGCGG	2169							
Db	3134	CATCTTTGAAGCACCCAGGGTGGCGG	3159							

RESULT 10
CFA347749
LOCUS CFA347749 3230 bp mRNA linear NAM 06-JAN-2003
DEFINITION Canis familiaris mRNA for p76RBE protein.
ACCESSION AJ347749
VERSION AJ347749.1 GI:23304859

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

p76RBE protein.
Canis familiaris (dog)
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1 Mircescu,H., Steuve,S., Savonet,V., Degraef,C., Mellor,H.,
Dumont,J.E., Maenhaut,C. and Pirson,I.
Identification and characterization of a novel activated RhoB
binding protein containing a PBZ domain whose expression is
specifically modulated in thyroid cells by cAMP
Eur. J. Biochem. 269 (24), 6241-6249 (2002)
12473120
2 (bases 1 to 3230)
Pirson,I.
Direct Submission
Submitted (20-SEP-2001) Pirson I., Interdisciplinary Institute
IRIBHN, Université Libre de Bruxelles, 1070, Brussels, BELGIUM
Location/Qualifiers
1. 3230
/organism="Canis familiaris"
/mol_type="mRNA"
/db_xref="taxon:9615"
22. .2082
/function="putative RhoB effector"
/codon_start=1
/product="p76RBE protein"
/protein_id="CAC87938.1"
/db_xref="GI:23304860"
/db_xref="GOA:Q8HXG3"
/translation="MTDALLPAQPOPLEKSDCYFRKGNPLQATGTRSKLQNRALN
QQILKAVRMRTGAENLLKVNTHKREVRLEISFVNSDLQMLKEELNLSVGYQ
STERDAPVLLPLGKQKIDFVSLKDFILEHYSEDSYLYEDLADMLDRLQACRT
PFRDAGVLLMSYFIQGFVSEFFPTPTQMGILFTWYDSLTPVPSQONLLKAS
ILFNIGALYTGITGCTNRRTAGLDGADVAFORAAAGVHLHKLKTTPTPSYDMSAML
SVLVKMLAQOENVEFKICLPGRNEFFVLKVAQAAKVEVVRQHLTAMSOAPVK
LTPYASLACVKAHYAAHAHVAATLLIDHQLKPGADEDHQKLSQLYDHMPG
LTELATKSGHQRLGKSHLRRAVHAHEESVREASLCKLNIEVLQDVLSVAHERS
RLKTAHQDDDDLLIDAPDIISKTEQVEIILPQFSKVTATDFPKGLPSVFSAN
KWTTPPSIHFTABEGDLFTLRGNSPVQVHFLDPHCASALAGAKEGDYIVSIQVDC
KLWTVSEYMKLLKACGDRGVEMKVVSLDFTSSMHNKATYSVGMOKTYSMCLAIID
DDKTDKTKKISKLSFLSWGTDKNRVKSATSLCLPSVGPVQVKKLPSPFLNSD
SSLY"

CDS

/function="putative RhoB effector"

/codon_start=1

/product="p76RBE protein"

/protein_id="CAC87938.1"

/db_xref="GI:23304860"

/db_xref="GOA:Q8HXG3"

/translation="MTDALLPAQPOPLEKSDCYFRKGNPLQATGTRSKLQNRALN
QQILKAVRMRTGAENLLKVNTHKREVRLEISFVNSDLQMLKEELNLSVGYQ
STERDAPVLLPLGKQKIDFVSLKDFILEHYSEDSYLYEDLADMLDRLQACRT
PFRDAGVLLMSYFIQGFVSEFFPTPTQMGILFTWYDSLTPVPSQONLLKAS
ILFNIGALYTGITGCTNRRTAGLDGADVAFORAAAGVHLHKLKTTPTPSYDMSAML
SVLVKMLAQOENVEFKICLPGRNEFFVLKVAQAAKVEVVRQHLTAMSOAPVK
LTPYASLACVKAHYAAHAHVAATLLIDHQLKPGADEDHQKLSQLYDHMPG
LTELATKSGHQRLGKSHLRRAVHAHEESVREASLCKLNIEVLQDVLSVAHERS
RLKTAHQDDDDLLIDAPDIISKTEQVEIILPQFSKVTATDFPKGLPSVFSAN
KWTTPPSIHFTABEGDLFTLRGNSPVQVHFLDPHCASALAGAKEGDYIVSIQVDC
KLWTVSEYMKLLKACGDRGVEMKVVSLDFTSSMHNKATYSVGMOKTYSMCLAIID
DDKTDKTKKISKLSFLSWGTDKNRVKSATSLCLPSVGPVQVKKLPSPFLNSD
SSLY"

ORIGIN

Query Match	25.8%;	Score 558.8;	DB 4;	Length 3230;
Best Local Similarity	59.0%;	Pred. No. 1.1e-77;		
Matches 1019;	Conservative	0;	Mismatches 697;	Indels 12; Gaps 3;
Qy	219	GCGCTGTGACTCCCTGACGCGCAGATCCAGTGGCGGCAGCTGCAGAGCCGAGGCCAGAT	278	
Db	90	GCGTTGCAATCCCTTTCGCAAACTGGCGGAGCAAACTGCAGAACTCAGAGGGCTGCCTT	149	
Qy	279	TCACGACAGATTGACAAAGGAGCTGCAGATCGGACGCGCGCTCAGAACCTCTACAGAGC	338	
Db	150	GAACACGACAGATCTCGAAAGCGCTGCGGATGAGAACGGAGCAAAACCTCTCGAAAGT	209	
Qy	339	CACGACAAACACCGGTGAGAGAGACGGTGGCCCTCGAGCTGAGCTACGTCACATCCAA	398	
Db	210	AGCTACGAACACCAAGGTAACGGAGAGGTGCGCTCGAGCTGAGCTTCGTGAACTCTGA	269	
Qy	399	CCTGCAGCTCTGAGGAGGAGCTGGAGGAGCTCAGCGGTGGCGGTGGACCCCTGGCGGCA	458	
Db	270	CCTGCAGATGCTCAAGGAAGAGCTGGAAAGGCTGGAACATCTCAGTGGAGTCTATCAGAG	329	
Qy	459	TGGAGCGAAAGTCTCACTGTCCCATGATCCCTCGGCTGAAGGAGACCAAGGAGCT	518	
Db	330	CACGAGGAGGCGTTTACGTTCCCTTGTATTCGCTTGGCTGGAAGAAACCAAGGACAT	389	
Qy	519	GGACTGCTTACACCGCTGAAGGAGCTGATCTCAGTGCACCTTTGGAGAGGACGCGCTC	578	

Db	390	TGACTTCTCAGTCGTTCTCAAGGATTTTCATCTCTGGAACATTACAGCGAGGACAGCTATCT	449
Qy	579	CTACGAGGAGAAATCAGGGAGCTTGAGGCGCTTGTGGGCGAGGCCCATGGGAGACCCCGAGCCG	638
Db	450	ATATGAAGACGAATAGCAGACCTGATGACCTGAGACAGGCTTGGCGGACGCCCGAGCCG	509
Qy	639	GAATGAGTCGGGCTGGAGTGTCTACAGGCTATTACACAGCTGTTCTCTGATGC	698
Db	510	GGATGGCGCGGGTTGAACCTGCTCATGAGCTACTTCATCAGCTGGGATTTGTGGAGAG	569
Qy	699	GGCCTTCTCCACCCCTGCCAGAGGCTTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC	758
Db	570	CCGCTTCTTCCCCCCCACCGACAGATGGGATCTTGTTTACCTGGTATGACTCCCTCAC	629
Qy	759	TGGGCTTCCGCGCCAGCAGCGTGCCTTGGGCTTCGAGAGGGCAGCGTCTCTTCAACAT	818
Db	630	CGGGCTCCCGGTGAGCCAGCAGAATTTGCTGCTGGAGAGGCCAGCATTTCTGTCAACAT	689
Qy	819	CGGTGCGCTCCACAGCAGATTTGGGCGCGCCAGGACCGCTCTGTCACCGAGGCTGCCG	878
Db	690	CGGAGCCCTCTACACAGATCGGGAACCCGGTGCATCGGGGACCCAACTGGGCTGGA	749
Qy	879	CCGCGCTATGAGGCTTCCAGAGGCGCGCTGGGGCTTTCAGCTCTCTGAGGAGAACTT	938
Db	750	CGGTGCGGTGAGCGCTTTCAGAGAGCGCGAGGGGTTTTCATCACCTGAAAGAGAGCTT	809
Qy	939	CTCCATGCGCGGAGCCAGCATGAGGCTGCGTCCCTCTGCGCATGCGCATGAGCAGCTCAT	998
Db	810	CACCTCACACTCCGAGTTACGACATGAGCCCTGCCATGCTCAGCGTGTGCTCAAGATGAT	869
Qy	999	GATGGCCAGGCCACAGAAATGTGTTTGAGGGCTCTCTCACCACTGCTCCATGGCCCC	1058
Db	870	GCTTGCACAGCCAGGAGACGTTGTGAGAAATCTGCTTCTG-----GGATCCG	923
Qy	1059	CCAAAGACTGCTGCGCCAGCTGCGCTGCGCGCAGAGGCGGCCAGGTGGGAGCCGAGTA	1118
Db	924	GAATGAGTCTTCTGCTGCTGTTGAAGTGTGCTCAGAGGCTGCCAAGGTGGGGAGGTGA	983
Qy	1119	CAGGCTAGTGACCGGACATGGCGCAGCACCCTGCTCCAGCACTACGTGCTGCTCTCTG	1178
Db	984	TCGGCAGCTGCACACGGCCATGAGCCAGGCCCGGTGAAGAGAACATCCCTTACTCTG	1043
Qy	1179	GACTGCTCTGTGATGTCGAAGGCGGAGTACTTCCGCTCCCTGCGCCACTACCACTGAGC	1238
Db	1044	GGCCAGCTGTGCTGCTGTAAGGCCACCACTAGCGGCCCTTGCCCACTACTTCCGAGC	1103
Qy	1239	CATGGCCCTCTGCGACGGCTCCCGACGACCGAGGAGAGTCCCGACGACGAGCGGT	1298
Db	1104	CACCTTCTCATCGACCAACCACTGTAAGCTGGTGCAGATGAGGACCATCAGAGAGATG	1163
Qy	1299	CTTCTGTGAGC-----CCCCGACTCTTAAGCCCGAGGCCCTGTGCTGCC--GCAGGA	1352
Db	1164	CCTGTCCAGCTCTACGACCACTATGCCAGAGGGACTGACGCCCTTGGCCACCCCTCAAG	1223
Qy	1353	GCTGGAGGAGCGCAGCAGCTTGGCAAGGCACACTCTGAAGCGTGTGCATCTCTGGGCGAGGA	1412
Db	1224	CGGCCACAGCGCCGACAGCTGGGCAAGTCCCACTGCGAGAGCGGTGGCCACCAAGA	1283
Qy	1413	GGAGCGCTGCGGCTGCAAGCCCTGTGCGCGGTCTGCGGAGGTGGACCTGCTTCGGGC	1472
Db	1284	GGAGTCGTGAGGGAGGGCCAGCCTGTGCAAGAAGCTCCGCAACATCGAGGTGCTCCAGGA	1343
Qy	1473	TGTGATCTCCAGACGCTGCAGCGCTCTAGTCCCAAGTATGCGGAGCTCGACCGTGAGGA	1532
Db	1344	CGTGTGTCCGCGGCGCAGAGCGGTCCCGCTCAAGTACGCGGACGACCGAGCAGGA	1403
Qy	1533	TGACTTCTGTGAGGCTGCCGAGGCCCGGACATCCAGCTTAAGACCCACCGAAGCCAGA	1592
Db	1404	CGATCTGTCTCACTTGATGATGCTCCCGACATATCTCTAAACTGAGCAAGAGGTTGA	1463
Qy	1593	GGCCAGGATGCCAGCGCTTCCAGGGGAGGGGCTTGACATCTTCCATCGCGCTGGGCC	1652
Db	1464	AATTATATGCCACAGTTCTTCCAAAGGTGACAGCAACGGGACTTCTTCCAGAAAGCTGGGCC	1523
Qy	1653	CCTGCTCTGTCTCTCAGCCAAAGAACCGGTGGCGGCTGGTGGGGCCCGTCCACCTGACCCG	1712
Db	1524	CCTGCTCAGTGTCTTTCGGCTTACAGAGATGGACACCTCCCTCGAAGCATTCACCTTCACTGC	1583
Qy	1713	AGGAGAGGGCGGCTTTGGCTTACCGCTTGGGAGAGACTGCGCTGTCTCTCATCGCTGCCGT	1772
Db	1584	AGAAGAAGGGATCTGGGGTTACCTTTGAGAGGAAACTTCCCCAGTGCAGTCCATTTCCT	1643
Qy	1773	CATTCCAGGAGCCAGCGCGCGCGCTGCGCTGAAGGAGGCGACTACATTTGTGTCACT	1832
Db	1644	GGATCTCTCACTGCTCTGCTGCGCTGGCGGAGCCAAAGGAGGATATATTTTTCAT	1703
Qy	1833	GAATGGCGCAGCCATGACAGGTGTGTGAGACACGCGGAGGTGTGACGAGCTGAAGGCTGC	1892
Db	1704	TCAGACGTGGATTGCAAGTGGCTGACGCTGAGCGAGGTCTATGAAGCTGCTGAAGGCGCTG	1763
Qy	1893	GGAGAGGGCGGGCGCCAGCGCTGACAGGTGTGTGCTGCTGCTGCCAGCTC	1940
Db	1764	TGGCAGGGACGCGCTGAGATGAAGGTGCTGAGCTCTCTGGACTTCAC	1811
RESULT 11			
BC036447			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
1 (bases 1 to 3516)			
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,			
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,			
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,			
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hele, F.,			
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,			
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,			
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,			
Armanaci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,			
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,			
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,			
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,			
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,			
Fahney, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,			
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,			
Boiffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,			
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,			
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,			
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.			
Generation and initial analysis of more than 15,000 full-length			
human and mouse cDNA sequences			
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)			
2388257			
12477932			
2 (bases 1 to 3516)			
Strausberg, R.			
Direct Submission			
Submitted (09-AUG-2002) National Institutes of Health, Mammalian			
Gene Collection (MGC), Cancer Genomics Office, National Cancer			
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,			
USA			
NIH-MGC Project URL: http://mgc.nci.nih.gov			
On Sep 16, 2003 this sequence version replaced gi:22209018.			
Contact: MGC help desk			
Email: cgapbs-remail.nih.gov			
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.			
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki			
Toshiyuki and Piero Carninci (RIKEN)			
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)			

DNA Sequencing by: Institute for Systems Biology
http://www.systemsbiology.org

contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 34 Row: j Column: 22
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 23510298.

FEATURES

source

1..3516

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:26769 IMAGE:4830913"

/tissue_type="Testis"

/clone_lib="NIH_MGC_97"

/lab_host="DH10B"

/note="Vector: pBluescript"

1..3516

/gene="RHPN2"

/note="synonym: p76RBE"

/db_xref="LocusID:85415"

37..2097

/product="rhophilin-like protein"

/protein_id="AAH36447.1"

/db_xref="GI:22209019"

/db_xref="LocusID:85415"

/translation="MTDALLPAQPQLEKNDYFRKGNPLAQTRSKLQNRALN
QQLKAVMRTEAENLLKVAATNSKREVRLEISFVNSDLQMLEELEGNTSVGVYQ
NTEAFTPLIPGUKETKQVDFVVLKDFILEHSEGYDEIDLMIDLRQACT
PSRDAGVELLMTYFIQGFVSRFPFTRQMLLFTWYDSLTPGVPSQOQLLEKAS
VLNFGALYQIGTRQOAGLESDAIDAFORAAVLNLYKDTFTHPSPDMSAPML
SVLWMLAQLACVKAHVAALAHYFTAILLDHQVKGPTDLDHQKLSQLYDHMPG
ENIPYASLASLACVKAHVAALAHYFTAILLDHQVKGPTDLDHQKLSQLYDHMPG
LTPATLNDQQRQLGKSHLRRAWHHEESVREASLCKKRSIEVLQKLVCAQERS
RUTYAQHQEDLLNIDAPSVAATKEVDIILDFQSKLITDTPFKLGLPVSFSAN
KWTTPSRIRFTEAGSDIGFTLRGNAPVQVHPFLDPCASVAGAREGDIYVSIQIVDC
KWLTLSEYMLKLSFGEDEIEKVVSLDSTSSMKNKATYSVMQKYSMTCLDID
DKTDKTKIKSKLSFLSWGNTKNQKQSASTILCLPSVGNARQVKKLPSFSLNSD
SSWY"

misc_feature

163..354

/note="Rhl; Region: Rhl repeat"

/db_xref="CDD:pfam02185"

misc_feature

367..825

/note="BRO1; Region: BRO1-like domain. This functionally uncharacterised domain is found in a number of signal transduction proteins, including Rhophilin and BRO1"

misc_feature

/db_xref="CDD:pfam03097"

1576..1809

/note="PDZ; Region: Domain present in PSD-95, Dlg, and ZO-1/2. Also called DHR (Dlg homologous region) or GLGF (relatively well conserved tetrapeptide in these domains). Some PDZs have been shown to bind C-terminal polypeptides".

ORIGIN

Query Match 22.9%; Score 496.4; DB 9; Length 3516;
Best Local Similarity 56.8%; Pred. No. 5.9e-68;
Matches 977; Conservative 0; Mismatches 731; Indels 12; Gaps 3;
Qy 219 GGGCTGTGACTCCCTGACGAGATCCAGTGGCGGCGAGTCCAGAGCCGCGAGCCAGGCGCCAGAT 278
Db 105 GGGCTGTGTAATCCCTTGCACAAACCGCGCGAGTAAATTCAGATCAAGAGCTGCTTT 164
Qy 279 TCACGAGCAGATTGACAGGAGCTGCAGATCGGACGGCGCTGAGAACCTCTACAGAGC 338
Db 165 GAATCAGCAGATCTCTGAAGCGGTGCGGATGAGGACCGGAGGAAACCTTCTGAAAGT 224
Qy 339 CACCAGCAACAACCGGTGAGAGAGACGGTCCGCTTGAGGCTGAGTCAAGTCAACTCCAA 398

Db 225 GGCCACAAACTCAAAAGTGGCGGAGCAAGTGGCGCTGGAGCTTGGTCACTCAGA 284
Qy 399 CTTGCGAGCTGCTGAAGGAGAGCTGGAGGAGCTAGCGGTGGCGTGGACCTGCGCCGCA 458
Db 285 CTTGCGAGTCTCAAGGAGAGCTGGAGGAGCTGGAACATCTCGGTGGCGCTCTATCAGAA 344
Qy 459 TGGGAGCGAAGCTGTCACTGCTCCCATGATCCCTCGGCTTGAAGGAGAGCCAGGAGCT 518
Db 345 CACAGAGGAGCATTTACGATTTCCCTGATCTCTTGGCTTGAAGGAAACGAAAGACGT 404
Qy 519 GGAATGCTTACACCGCTGAGGAGCTGATCTCAGTGCACTTTGGAGAGAGCGCGCTC 578
Db 405 CGAATTTGCGAGTCTCTCAAGGATTTATCTCTGGAACATTAAGTGAAGATGGCTATTT 464
Qy 579 CTACGAGGAGCAAAATCAGGAGCTGGAGGCTTGGCGGAGGCGCATGCGGACCCCGAGCG 638
Db 465 ATATGAAGATGAAATTTGCAGATCTTATGATCTGAGACAAGCTTGTGCGACGCTAGCG 524
Qy 639 GAATGAGTGGGCTTGGAGCTGCTCACAGCTTATTAACACCACTGTGCTTCTGGATGC 698
Db 525 GGATGAGGCGGGTGGAACTGCTGATGACATATCTTATCCAGCTGGGCTTTGTCGAGAG 584
Qy 699 GGGCTTCTCACCCCTGCGAGGAGCTCGGGCTCTTTTCCACTGGTACGACTCGCTTAC 758
Db 585 TCGATTTCTCCGCGCCACACGCGAGATGGAGCTCTCTGTTTACCTGTATGACTCTCTCAC 644
Qy 759 TGGGGTCCCGGCGGAGGAGCTGCGCTTGGAGGAGGAGGAGGCTTCTCTTCAACAT 818
Db 645 CGGGGTTCGGGTACGACAGAACTGCTGTGGAGAAAGGCGAGTGTCTCTGTCAACAC 704
Qy 819 CGGTGCTCTCCACACGAGATTTGGGCGCGCGAGGACCGCTCTCTGACCGAGGTCGCCG 878
Db 705 TGGGGCTCTTACACCCAGATTTGGGACCGGTGTGATCGGAGACCGAGGCTGGGCTGGA 764
Qy 879 CGCGCTATGAGGCTTCCACAGGCGCGCTGGGCGCTTTCAGCGCTCTCTGAGGAGAACTT 938
Db 765 GAGTGCCATAGATGCTTTTACAGAGCGCGAGCGGGGTTTAAATACCTGAAAGACATTT 824
Qy 939 CTCCTCATCGGCGGCGGAGAGCATGAGCGCTGCTGCTCTCTGCGCACTGGAGAGCTCAT 998
Db 825 TACCATATCTCAAGTTACGACATGAGCGCTGCCATGCTCAGCGTGTCTGCTCAAAATGAT 884
Qy 999 GATGCGCCAGGCGCCAGGAATGTGTGTTGAGGCGCTCTCACCACTGCTCTCCATGCCCC 1058
Db 885 GCTTGCACAGCCCAAGAAAGCGTGTGTTGAGAAATCAGCCTTCTCTG-----GGATCG 938
Qy 1059 CAAGACTGCTGGCGCCAGCTGCGCTGGCGAGAGGCGCGCCAGGTCGCGAGCGAGTA 1118
Db 939 GAATGAATTTCTTCACTGCTGGTGAAGTGGCTCAGGAGGCTGCTAAGTGGGAGAGTCTA 998
Qy 1119 CAGGTAGTGCACCGGACCATGCGCCAGCGCACCGGTCCAGCATGCTGCTGCTGCTCTG 1178
Db 999 CAACAGCTACCGGACCATGAGCGCGCGGTGAAGAGAAACATCCCTACTCTCTG 1058
Qy 1179 GACTGCTCTGCTGATGCTCAAGGCGGAGTACTTCCGCTCTCTGCGCCACTACCGAGTAC 1238
Db 1059 GCGCAGCTTAGCTGCTGAGGAGGCGCCACCATAGCGGCTTGGCCCTTGGCCCACTCAAGAA 1118
Qy 1239 CATGGCTCTTGGAGCGGCTCCCGAGCGAGCCGAGGAGAGTCTCCCGACGAGCAGGT 1298
Db 1119 CATCTCTCTCATCGACCAACAGGTGAAGCCAGGACCGGATCTGGACCAACAGGAGAGT 1178
Qy 1299 CTCTCTGAGC-----CCCCCAGCTCTTAAAGCCCGAGGCGCTGTGTC--TGCCGAGGA 1352
Db 1179 CCTGTCCAGCTCTTACGACCATGCGCAGAGGGGCTGACACCTTGGCCCACTCAAGAA 1238
Qy 1353 GCTGGAGGAGCGGAGGAGCTTGGCAAGGACACCTGAAAGCGTCCCATCTCTGGGCGAGGA 1412
Db 1239 TGATCAGAGCGGCGAGCTGAGGAGTCCCATCTTGGCAGAGCCATGCTCATACGA 1298
Qy 1413 GGAGGCGCTGCGGTGACCGCTCTGCGCGAGGTGACCTGCTTCCGGC 1472

Db 1299 GGAGTCGGTGGGAGCCAGCCTCTGCAAGAGAGCTGGGAGCAATTGAGTGTCTACAGAA 1358
QY 1473 TGTGATCTCCAGACGCTGACGGCTCACTGGCCAAAGTATGCGAGCTGACCGTGAGGA 1532
Db 1359 GGTGCTGTGGCCGACAGGAACGCTCCCGCTCACTGACGCCAGCACCAGGAGGGA 1418
QY 1533 TGACTTCTGTGAGGCTGCGAGGCGCCCGGACATCCAGCCTAAGACCCAGAGCCAGA 1592
Db 1419 TGACCTGCTGAACCTGATCGACGCCGCCAGTGTGTGCTTAAACTGAGCAAGAGTTGA 1478
QY 1593 GGCAGAGATCCAGCCTGTCCAGAGGAGGCGCTGACATCTTCATCGCTGGGGCC 1652
Db 1479 CATTATATTGCCAGTCTCCAGCTGACAGTCAGTCAGGACTTCTTCAGAGCTGGGCC 1538
QY 1653 CTTGCTGTGTCTCAGCCAAAGAACCGTGGCGGCTGGTGGGCGCCCTCCACCTGACCCG 1712
Db 1539 CTTATCTGTGTTTTCGGCTAACAAGCGGTGGAAGCGCTCCCTCGAAGCATCCGCTTCACTGC 1598
QY 1713 AGAGAGGGCGGCTTTGGCTCAGCTTCGGGAGAGCTCGCCTGCTCTCTCATCGTGCCT 1772
Db 1599 AGAAGAGGGGACTTGGGGTTTCACTTTGAGAGGGAACGCCCGCTTCAGGTTCACTTCT 1658
QY 1773 CATTCCAGGAGGAGCGCGCGCTGGCTGAAGGAGGCGGACTACATTGTGTGAGT 1832
Db 1659 GATCCTTACTGCTCTGCTCGGTGGCAGGAGCCCGGAGGAGATTATTTGCTCCAT 1718
QY 1833 GAATGGGACGCAATGCAAGGTGGTGAGACAGCGGAGGTTGTGACGAGCTGAAGGCTGC 1892
Db 1719 TCAGCTGTGATTTGAATGAGTGGCTGACGCTGAGTGAGGTTATGAAGCTGCTGAAGAGCTT 1778
QY 1893 GGGAGAGCGGCGCGAGCTGCGAGTGGTGTGCTGCTGCTG 1932
Db 1779 TGGCGAGGACGAGATCGAGATGAAGTGTGAGGCTCTG 1818

RESULT 12
HSA347750
LOCUS HSA347750 2061 bp mRNA linear PRI 06-JAN-2003
DEFINITION Homo sapiens mRNA for p76RBE protein.
ACCESSION AJ347750.1
VERSION AJ347750.1 GI:23304861
KEYWORDS p76RBE protein.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
Mircescu,H., Steuve,S., Savonet,V., Degraef,C., Mellor,H.,
Dumont,J.E., Maenhaut,C. and Pirson,I.
Identification and characterization of a novel activated RhoB
binding protein containing a PDZ domain whose expression is
specifically modulated in thyroid cells by cAMP
Eur. J. Biochem. 269 (24), 6241-6249 (2002)
23361238
12473120
2 (bases 1 to 2061)
Pirson,I.
Direct Submission
Submitted (20-SEP-2001) Pirson I., Interdisciplinary Institute
IRIBHN, Université Libre de Bruxelles, 1070, Brussels, BELGIUM
FEATURES
source
1..2061
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..2061
/function="putative RhoB effector"
/codon_start=1
/product="p76RBE protein"
/protein_id="CAC87939.1"
/db_xref="GI:23304862"
/db_xref="GOA:Q8IUC4"
/db_xref="SPTREMBL:Q8IUC4"

CDS

/translation="MTDALLPAAPQPLEKENDGYFRKGNPLAQTGRSKLQNRALIN
QOILKAVRVTGAENLLKVATNSKVRQVLELSFVNSDLQMLKELEGLNLSVGYQ
NTEETIPLIPLIGKETKDVFAVLKDFILHEHYSEDGLVLEDELDLMDLQACRT
PRDEAGVELLWTFYQLGFVSRPEPTROMGLLATWVDSLTPVPSOONLLLEKAS
VLFNTALYTGICRDRQTQAGLESADAFQRAAGVLYNLKDTFTFTPSYDMSAML
SVLTKMVAQAQESVEFKISLPIRNEFFMLVKVQAQAAGVGYOOLHAMSQAQPK
ENIPYSWASLACVKAHYAALAHYFTAILLIDHVPKPGTDLDOEKLSQLYDHMEPG
LTPLATLKNDOOROLKSHLRAMAHESVRELSCKLRSIEYLOKVLCAQERS
RLTYAOHEDDLNLLDAPSVAKTEQVDIILPOFSKLTVTFQKLGPLSVFSAN
KWTTPRSIRFRAEGDGLFTLGNAPVOVHFLDPYCSASVAGAREGDTVSLQVDC
KMLTSEVWLLKLSFGDEBIEMKVSVLLSTSSMHNKSATYSVGMQKTYSMICLAIDD
DKTDTKTKISKLSFLSGTWNKROKASTLCLPSVGAARPQVKKLPSFFLLNSD
SSWY"

ORIGIN

Query Match 22.8%; Score 494.8; DB 9; Length 2061;
Best Local Similarity 56.7%; Pred. No. 1.2e-67;
Matches 976; Conservative 0; Mismatches 732; Indels 12; Gaps 3;
QY 219 GGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGCGAGCTGCAGAGCGCGAGGCCAGAT 278
Db 69 GGGCTGTATCCCTTTCACAAACCGCGAGTAAATTCAGAAATCAAAGAGCTGCTTT 128
QY 279 TCACCAGCAGATTGACAAAGGAGCTGCAGATGCGGAGCGGCTGAGAACTCTACAGAGC 338
Db 129 GAATCAGCAGATCTCGAAAGCCGTGGCGATGAGACCGGAGCGGAAACCTTCTGAAAGT 189
QY 339 CACCAGCAACACCGGCTGAGAGACGCTGCGCTGGAGCTGAGCTACCTCAACTCCAA 398
Db 189 GGGCACAACCTCAAAAGTGGGGAGCAAGTGGCGCTGGAGCTGAGCTTCGTCAACTCAGA 248
QY 399 CCTGCAGCTGCTGAAGGAGGAGCTGGAGGAGCTACGCGTGGGCTGAGCCCTGGCGCGCA 458
Db 249 CTTGCAGATGCTCAAGGAGAGCTGGAGGGGCTGAACATCTCGGTGGCGCTCTATCAGAA 308
QY 459 TGGGAGCGAAGCTGCTGCTGCCATGATCCCTCGGCTGGAGCTGAGGAGCCAGAGGAGCT 518
Db 309 CACAGAGGAGGCAATTTACGATTCCTCTTGGCGCTGAAGAAACGAAAGAGCT 368
QY 519 GGACTGCTTACACCGCTGAAGGAGCTGATCTCAGTGCATCTTGGAGAGGAGCGCGCTC 578
Db 369 CGACTTTGCGAGTCTCTCAAGGATTTATCTCGAACATTTACAGTGAAGATGCTATTT 428
QY 579 CTACGAGCGAGAAATCAGGAGCTGGAGGCGCTCGGCGAGGCCATGGAGACCCCGAGCG 638
Db 429 ATATGAAGATGAATTTGAGATCTTATGATCTGAGACAAAGCTTGTGCGAGCGCTAGCG 488
QY 639 GAATGAGTGGGCGCTGGAGCTGCTCAGAGCTATTACACAGCTGTGCTTCTCTGGATGC 698
Db 489 GGATGAGCGCGGGTGGAACTGCTGATGACATCTTATCAGCTGGGCTTTGTGCGAG 548
QY 699 GCGCTTCTCACCCTCGCAGGAGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC 758
Db 549 TCGATTTCTCCCGCCACACGCGAGATGGGACTCTGTTCACCTGATGACTCTCTCAC 608
QY 759 TGGGGTCCCGGCCCGCAGCAGCGTGCCTTGGAGGAGGCGAGCGTCTCTTCAACAT 818
Db 609 CGGGTTCCGGTTCAGCCAGCAGAGAACCTGCTGCTGGAGAGGCCAGCTGTCTGTTCACAC 668
QY 819 CGGTGCGCTTCCACAGCAGATTTGGGCGCGCAGAGCCGCTCTCTCACCGAGGGTGC 878
Db 669 TGGGGCGCTTACACCCAGATTGGGACCCCGGTGTGATCGGAGAGCGAGGCTGGGCTGA 728
QY 879 CCGCGCTATGGAGGCGCTTCCAGAGGCGCGCTGGGCGCTTTCAGCGCTCTCGAGGGAGAACTT 938
Db 729 GAGTGCCATGATGCTTTTCAGAGAGCGCGAGGGGCTTTTAAATTAACCTGAAAGACACAT 788
QY 939 CTCCCATGCGCGAGCCAGACATGAGCGCTGCGTCCCTCTCTCGGACCTGGAGCGACTCAT 998
Db 789 TACCCATATCTCAAGTTTACGACATGAGCCCTGCCATGCTGAGCGCTGCTGCTCAAAATGAT 848
QY 999 GATGCGCCAGCGCCAGGATGTGCTTTGAGGGCGCTCTCACACCTGCTCTCATGGCCCC 1058

```
Db 849 GCTTGACAAAGCCCAAGAAAGCGTGTGAGAAATCAGCCTTCCTG-----GGATCCG 902
QY 1059 CCAAGACTGCTGGCCAGCAGTGGCCCTGGCGAGAGGCGCCGACAGTGGCAGCCGAGTA 1118
Db 903 GAATGAATTTCTCATGCTGGTGAAGGTGGCTCAGAGGCTGCTAAAGTGGGAGAGGTCTA 962
QY 1119 CAGGCTAGTGCACGCGACCATGCGCCAGCCAGCCCGTCCAGCACTAGTGTCTCTCTG 1178
Db 963 CCAACAGCTACAGCGACCATGAGCCAGCGCGCGGTGAAAGAAACATCCCTACTCTCTG 1022
QY 1179 GACTGCCCTGTGTGATGTCAAGGCGGAGTACTTCCGCTCCCTGGCCCACTACCACTAGC 1238
Db 1023 GGCACAGCTTAGCTGTGAGGCGCCACCACTAGCGGCGCTGGCCCACTACTTCACTGC 1082
QY 1239 CATGGCCCTCTGGAGCGGTCCCGACGACCGAGGAGAGCTCCCAOAGCAGCAGGT 1298
Db 1083 CATCTCTCTCATGACCAACAGGTGAAGCCAGGACCGGATCTGGACCAACAGGAGAAGTG 1142
QY 1299 CTTCTCTGCAGC---CCCCACCTCTCTAAGCCCGAGGCGCTGTGC--TGCAGCAGGA 1352
Db 1143 CTTCTCTCTCATGACCAACAGGTGAAGCCAGGACCGGATCTGGACCAACAGGAGAAGTG 1142
QY 1353 GCTGAGGAGCGCAGCAGCTTGGCAAGGCACACCTGAAGCGTGCCTCTCTGGGCGAGGA 1412
Db 1203 TGATCAGCAGCGCGAGCAGCTGGGAGTCCCACTTGGCAGAGCCATGGCTCATCAGA 1262
QY 1413 GAGGCGCTGCGGCTGCACGCTGTGCGCGGTCTTGGCGAGGTGGACCTGCTTCCGGCG 1472
Db 1263 GAGTGGGTGGCGGAGGCGAGCTCTGCAAGAGCTGCGAGCATTGAGGTGCTACAGAA 1322
QY 1473 TGTGATCTCCAGAGCGCTCAGCGCTCACTGCGCAAGTGAAGCGTGCCTCTCTGGGCGAGGA 1532
Db 1323 GGTGCTGTGTGCGCACAGGAAGCTTCCCGGTTCAGTACGCGCCAGCACCCAGGAGGAGGA 1382
QY 1533 TGACTTCTGTGAGGTGCGAGGCGCCGAGCATCCAGCCTAAGACCCACCAAGAGCCAGA 1592
Db 1383 TGACCTGTGAACCTGATCGACGCGCCCGAGTGTGTGCTTAAACTGAGCAAGAGTTGA 1442
QY 1593 GGCAGAGTGCACGCTGTCCAGGGGAAGGGGCTTGACATCTTCCATTCGCTGGGCGC 1652
Db 1443 CATTAATATGCCCCAGTTCTCCAAGCTGACATCAGGACTTCTTCCAGAGCTGGGCGC 1502
QY 1653 CTTGTCTGTGTCTCAGCAAGAACCGGTGGCGGCTGTGGGCGCGTCCCTCAAGCATCCGCTTCACTGC 1712
Db 1503 CTTATCTGTGTTTCGGCTTACAGCGGTGAGCGCTCTCTCAAGCATCCGCTTCACTGC 1562
QY 1713 AGGAGGCGCGCTTGGCTCAGCGCTTCCGGGAGACTCGCTGTCTCATCGCTGCGGT 1772
Db 1563 AGAAGAGGGGACTTGGGGTTTCACTTGAGAGGGAACGCCCGTTCAGGTTCACTTCT 1622
QY 1773 CATTCAGGAGCCAGCGCGCGGCTGGCTGAGGAGGCGGCTACATTTGTCTCAGT 1832
Db 1623 GGATCTTACTGCTGCTGCTCGGTGGCAGGAGCCCGGGAAGAGATTAATTTGTCTCCAT 1682
QY 1833 GAATGGGCGAGCATCAGCTGTGTGAGACACCGGAGGTGTGTGAGGAGCTGAAGGCTGC 1892
Db 1683 TCAGCTTGTGATTGTAGTGTGCTGACGTGAGTGGTGTGAGGCTGTGAAGGCTT 1742
QY 1893 GGGAGAGCGGCGCGCAGCTGCAAGTGTGTGCTGCTGCTG 1932
Db 1743 TGGCAGGACGAGATCGAGATGAAGTGTGTGAGGCTCTG 1782
```

RESULT 13

```
AX835273 LOCUS 3500 bp DNA linear PAT 15-DEC-2003
DEFINITION Sequence 2397 from Patent EP1347046.
ACCESSION AX835273
VERSION AX835273.1 GI:39921408
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
REFERENCE
AUTHORS Iisogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
Yamanoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and
Masuhio,Y.
TITLE Full-length cDNA sequences
JOURNAL Patent: EP 1347046-A 2397 24-SEP-2003;
RESEARCH Research Association for Biotechnology (JP)
FEATURES
source Location/Qualifiers
1..3500
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
```

```
ORIGIN
Query Match 22.8%; Score 494.8; DB 6; Length 3500;
Best Local Similarity 56.7%; Pred. No. 1e-67;
Matches 976; Conservative 0; Mismatches 732; Indels 12; Gaps 3;
QY 219 GGGCTGTGACTCCCTGACGAGATCCAGTGGCGGAGCTGCAGAGCGGCGAGGCCAGAT 278
Db 105 GGGCTGTATCCCTTGCACAAACCCGCGGAGTAAATTGCAGAAATCAAGAGCTCTTT 164
QY 279 TCACGACGAGATTGACAGGAGCTGCAGATGCGGACGGCGCTGAGAACTCTACAGAGC 338
Db 165 GAATCAGCAGATCTCGAAGCCATCGGATGAGGACCGGAGCGGAAACCTTCTGAAAT 224
QY 339 CACCGACAAACCGGCTGAGAGAGCGGTGCGCTTGGAGCTGAGCTCAACTCCAA 398
Db 225 GCGCAAAACTCAAGGTGCGGAGCAAGTGGCGGTGAGCTGAGCTTCTGCAACTCAGA 284
QY 399 CTGCAAGCTGTGAAGGAGGAGCTGAGAGAGCTCAGCGTGGCGTGGACCTTGGCGGCA 458
Db 285 CTTGAGATGCTCAAGAGAGAGCTGGAGGGCTGAACATCTCGTGGGCTCTATCAGAA 344
QY 459 TGGGAGCGAAGCTGTCACTGTCCCATGATCCCCCTTGGGCTTGAAGGAGCAAGAGCT 518
Db 345 CACAGAGGAGGATTTACGATTTCCCTGATTTCCCTTGGCTTGAAGGAAACGAAAGCT 404
QY 519 GGAAGTGTCTACACCGCTGAGGAGCTGATCTCAGTGCACTTTGGAGAGGACGGCGCTC 578
Db 405 CGACTTTGAGTGTCTCAAGGATTTTATCTGGAACATTTACAGTGAAGATGGCTATTT 464
QY 579 CTACGAGCAGAAATCAGGAGCTGAGCGCTGCGGAGGAGGAGGAGGAGGAGGAGGAG 638
Db 465 ATATGAAGATGAATGAGATCTTATGATCTGAGACAAGCTTGTGCGAGCGCTTAGCGG 524
QY 639 GAATGAGTGGGCTTGGAGCTGCTCACAGCTTATTTACAAACAGCTGTGCTTCTCGATGC 698
Db 525 GATGAGGCGCGGGTGGAACTGCTGATGACATCTTCACTCAGCTGGGCTTGTGAGAG 584
QY 699 GGGCTTCTCAACCTTGCAGAGGCTCGGGCTCTTTCACCTGGTACGACTGGCTTAC 758
Db 585 TCGATTTCTCCGCGCCACACGCGAGATGGGACTCTCTTCACTGGTATGACTCTCTCAC 644
QY 759 TGGGCTCCCGCCCGCAGAGCGTGGCGCTTTCGAGAGGAGGAGGAGGAGGAGGAGGAG 818
Db 645 CGGGTTCCGCTCAGCCAGCAAGAACCTGCTGCTGGAGAGGAGGAGGAGGAGGAGGAG 704
QY 819 CGGTGCCCTTCCACACGAGATTTGGGCGCGCAGAGCGCTTCTGCAACGAGGAGGAGG 878
Db 705 TGGGCGCTCTACACCCAGATTGGGACCCCGTGTGATCGGAGAGGAGGAGGAGGAGGAG 764
QY 879 CCGGCTATGAGAGGCTTTCAGAGGCGCTTGGGCGCTTTCAGCTTCTGAGGAGAGACT 938
Db 765 GAGTGCCTATGATGCTTTCAGAGAGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 824
QY 939 CTCCATGCGCGGAGCCAGACATGAGCGCTGCTCCCTCTGCGCACTGGAGGAGGAGCT 998
Db 825 TACCATCTCCAAAGTTACGACATGAGCGCTTGGCAATGCTCAGCGTGTCTGCTCAAAAT 884
QY 999 GATGGCCCGCGCCAGGAATGTGTGTTGAGGGCTCTCTCACCACCTGCTCCTCATGGCCC 1058
```


Db	885	GCTGACAGCCAGCAAGAGCGGTGTGAGAAAATCAGCTTCTCTG-----GGATCCG	938	
QY	1059	CCAAAGACTGCTGGCCAGCTGCGCTTGGCGCAGGAGCGCCAGGTGGCAGCCGAGTA	1118	
Db	939	GAATGAATCTTTCATGCTGTGAAGTGGCTCAGGAGGCTGCTAAGGTGGAGAGGTCTA	998	
QY	1119	CAGGCTAGTGCACCGGACCAATGGCCAGCCACCCGCTCCACGACTACGTGCTGTCTCTG	1178	
Db	999	CCAAACAGCTACACGACGACCATGAGCCAGGCGCGGTGAAGAGAAATCATCCCTACTCTCTG	1058	
QY	1179	GACTGCGCTGGTGCATGTCAAGCGCGAGTACTTCGCTCCCTGCGCCACTACACAGTGAGC	1238	
Db	1059	GGCCAGCTTAGCTGCTGAGGCGCCACCACTACGCGGCCCTGGCCCACTTCTACTGC	1118	
QY	1239	CATGGCCCTCTGGACCGCTCCCGACGCGCCGAGGAGAGCTCCCCACGACACGACGAGGT	1298	
Db	1119	CATCTCTCTCATCGACCAACGAGTGGAAGCCAGGACGCGATCTGGACCAACGAGGAAGTG	1178	
QY	1299	CTTCTCTGACG-----CCCCACACTCTCTTAAGCCCCGAGGCCCTGTGC--TGCGCGAGGA	1352	
Db	1179	CTGTCTCCAGCTTACGACCACTATCCAGAGGGGCTGACACCCCTTGCCCACTCTGAAGAA	1238	
QY	1353	GCTGGAGGCGGAGCGAGCTTGGCAAGGCACACCTCGAAGCGTGCCTATCTGGGGCAGGA	1412	
Db	1239	TGATCAGCAGCGCCGACAGCTGGGGAAGTCCCACTTGGCAGAGCCATGGCTCATCAGA	1298	
QY	1413	GGAGCGCTGCGGTGACGCGCTGTGCGCGCTCTCGCGAGGTGAACCTGCTTCGGGC	1472	
Db	1299	GGATCGTGGGAGGCGAGCGCTCTGCAAGAGCTGCGGAGCATTTAGGTGCTACAGAA	1358	
QY	1473	TGTGATCTCCAGACGCTGACGCTCACTGGCCCAAGTATGCGAGCTCGACCTGAGGA	1532	
Db	1359	GGTCTGTGTGCGCCACAGGAACGCTCCCGCTCACGTACGCTAGCCCGCAGCACGAGGAGGA	1418	
QY	1533	TGACTCTGTGAGGCTGCGGAGGCCCGGACATCAGCCTAAGACCCACCAAGAGCCAGA	1592	
Db	1419	TGACCTGTGAACTGTATCGAGCGCCCGCCAGTGTGTGCTAATAACTGAGCAAGAGTTGA	1478	
QY	1593	GGCAGGATGCCAGCTGTCCCGAGGGAAGGGCCGTGACATCTTCATCGGCTGGGGCC	1652	
Db	1479	CATTATATGCCCCAGTTCTCCAAGCTGACAGTCACGAGCTTCTTCCAGAAAGCTGGGCC	1538	
QY	1653	CCTGTCTGTGTTCTCAGCAAGAACCGGTGCGGCTGTGGGGCCCGCTCCACCTGACCG	1712	
Db	1539	CTTATCTGTGTGTTTGGGTAAAGCGGTGAGCGCTCTCTCGAAGCATCCCGCTTCACTGC	1598	
QY	1713	AGGAGAGGGCGCTTTGGCTTCAAGCTTTCGGGAGACTCGCTGTCTCTCATCGCTGCCGT	1772	
Db	1599	AGAAGAAGGGACTTTGGGGTTTCACTTGAGAGGGAACGCCCGCTTCAAGTTCACTTCT	1658	
QY	1773	CATTCCAGGGACCCAGCGCGGCTGGCTGAGAGGGCGGACTCATTTGTGTGAGT	1832	
Db	1659	GGATCTTACTGCTCTGCTCGGTGGCAGGAGCCCGGGAAGAGATTTATTTGTCTCCAT	1718	
QY	1833	GAATGGGAGCCATGCACTGTGTGAGACACCGCGAGGTGTGAGGAGCTGAGGCTGC	1892	
Db	1719	TCAGCTTGTGATTGTAAGTGGCTACGCTGAGTGAGGTTATGAAGCTGCTGAAGAGCTT	1778	
QY	1893	GGGAGAGCGGCGCGCCAGCTGACGCTGCTGCTGTGCTGCTGCTG 1932		
Db	1779	TGGCGAGGACGAGATCGAGATGAAGTGTGAGCTTCTCTG 1818		
RESULT 14	AK098246	3500 bp	mRNA	linear
LOCUS	AK098246			
DEFINITION	Homo sapiens cDNA FLJ40927 fis, clone UTERU2006593, weakly similar to GTP-RHO BINDING PROTEIN 1.			
ACCESSION	AK098246			
VERSION	AK098246.1			
KEYWORDS	oligo capping; fis (full insert sequence).			
SOURCE	Homo sapiens (human)			

ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 Oshima,A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K., Arita,M., Musashino,K., Yuuki,H., Harada,H., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Negahari,K., Masuno,Y., Nagai,K. and Isogai,T.
TITLE	NEDO human cDNA sequencing project
JOURNAL	Unpublished
AUTHORS	2 (bases 1 to 3500)
TITLE	Isogai,T. and Yamamoto,J.
JOURNAL	Direct Submission
COMMENT	Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB.
FEATURES	location: HRI and RAB.
source	1. 3500
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="UTERU2006593"
	/tissue_type="uterus"
	/clone_lib="UTERU2"
	/note="cloning vector: pME18SFL3"

Query Match	22.8%;	Score	494.8;	DB	9;	Length	3500;
Best Local Similarity	56.7%;	Pred. No. le-67;					
Matches	976;	Conservative	0;	Mismatches	732;	Indels	12;
				Gaps			3;

QY	219	GGGCTGTGACTCCCTGACGAGATCCAGTGGCGGAGCTGAGAGCCGAGCCGAGCCGAGAT	278	
Db	105	GGGCTGTGATCCCTTGCACAAACCGCGCGAGTAAATTGAGAATCAAGAGCTGCTTT	164	
QY	279	TCACGAGAGATTGACAGAGCTGAGATGCGGACGCGGCGCTGAGACCTTACAGAGC	338	
Db	165	GAATCAGCAGATCTGAAAGCCATCGGATGAGGACCGGAGCGGAAACCTTCTGAAAGT	224	
QY	339	CACGAGCAACAACCGGTGAGAGAGCGTCCCTGGAGCTGAGCTACGCTCAACTCAA	398	
Db	225	GGCCACAAACTCAAAAGTGGCGGAGCAAGTGGCGCTGAGCTGAGCTTCTCACTCAGA	284	
QY	399	CCTGAGCTGTGAAGGAGGAGCTGGAGAGCTCAGCGGTGGCGTGGAGCCCTGGCGCGCA	458	
Db	285	CCTGAGATGCTCAAGGAGAGCTGGAGGGCTGAACATCTCGGTGGCGCTTATCAGAA	344	
QY	459	TGGGAGCGAAGCTGTCACTGTCCCATGATCCCCCTGGCGCTGAGGAGACGAGGAGCT	518	
Db	345	CACAGAGGAGGCAATTTACGATTCCCTGATTTCCCTTTGGCGCTGAGGAGAAAGAAAGAGT	404	
QY	519	GGACTGTGTACACCGCTGAAGGAGCTGATCTCAGTGCATTTTGGAGAGGAGCGCGCTTC	578	
Db	405	CGACTTTCAGTCTGCTCTCAGGATTTTATCTTGAAATTTACAGTGAAGATGGCTATTT	464	
QY	579	CTACGAGCAGAAATCAGGAGCTGGAGCCCTGGCGGAGGCGCATCGGACCCCGAGCGG	638	
Db	465	ATATGAAGATGAATTGAGATCTTATGGATCTGAGACAAGCTTGTTCGAGACGCTTAGCGG	524	
QY	639	GAATGAGTGGCGCTGGAGCTGCACAGCCCTATTACAAACAGAGCTGTCTCTCGATGC	698	

Db 525 GGATGAGCGCGGGTGGAACTGCTGATGACATATTCATCCAGCTGGGCTTTGTGAGAG 584
Qy 699 GCGCTTCTCACCCTCGCAGGAGCCTCGGGCTCTTCTTCCATCTGTTGACACTCGCTTAC 758
Db 585 TCGAATCTTCCGCGCCACACGCGAGATGGGACTCCTGTTACCTGTTATGACTCTCTCAC 644
Qy 759 TGGGGTCCCGGCCAGCAGAGCTGCTCGGCTTCGAGAAAGGAGAGCTTCTTCAACAT 818
Db 645 CGGGGTTCGGTCAAGCAGAGAACTGCTGCTGGAGAGGCGAGTCTCTGTTTCAACAC 704
Qy 819 CGGTGCCCTCCACACGAGATTGGGGCGCGCCAGACCGCTCCTGACCGAGGCTGCCG 878
Db 705 TGGGGCCCTCTACACCCAGATTGGGACCCCGTGTGATGCGCAGCAGCGCTGGCTGGA 764
Qy 879 CCGGCTATGGAGGCTTTCAGAGGGCGCTGGGGCCCTTCAAGCTCTCTGAGGAGAACTT 938
Db 765 GAGTGGCATAGATGCTCTTTCAGAGAGCGCGAGGGGTTTAAATTAACCTGAAGACACATT 824
Qy 939 CTCCCATGGCCGAGCCAGACATGAGCGCTGCTCCCTGCGGCACTGAGCAGCTCAT 998
Db 825 TACCCATACTCCAAAGTTACGACATGAGCCCTGCCATGCTCAGCGTCTCGTCAAAATGAT 884
Qy 999 GATGGCCCGCCGACCAAGATGTGTGTTTGGAGGCTCTCACACCTGCGCTCCATGGCCCC 1058
Db 885 GCTTGACAAAGCCCAAGAAAGCTGTTTGAAGAAATCAGCTTCTCTG-----GATCCG 938
Qy 1059 CAGAGCTGCTGGCCAGCTGGCCCTGGCGAGGAGCGGCCAGAGTGGCAGCCGAGTA 1118
Db 939 GAATGAATTTCTTCATGCTGTGTAAGTGGCTCAGAGAGGCTGCTAAGTGGGAGAGTCTA 998
Qy 1119 CAGGCTAGTGACCGGACCATGSCCAGACCCCGCTCCAGCTAGCTAGTGGCTGTCTCTG 1178
Db 999 CCAACAGCTACACGAGCATGAGCAGGCGCGCGTGAAGAGAAATCCCTCTACTCTG 1058
Qy 1179 GACTGCCCTGGTGCATGTCAAGGCCAGTACTTCCGCTCCCTGGCCCACTACCACTAGC 1238
Db 1059 GGCAGCTTAGCTGCTGAAGSCCACCACCTAGCGGCGCTGGCCCACTACTTCACTGC 1118
Qy 1239 CATGGCCCTTGGAGCGCTCCCGACGACCGAGGAGAGTCCCCACGACGAGCAGGT 1298
Db 1119 CATCTCTCTCATGACACCAAGTGAAGCCAGGCAAGGATCTGGACCAACGAGAGAGTG 1178
Qy 1299 CTTCTCTGACG-----CCCCACCTCTCTTAAGCCCGAGGCGCTGTC--TGCCGAGGA 1352
Db 1179 CTGTCTCCAGCTCTAGCACCATGCGAGGAGGCTGACACCCCTTGGCCACACTGAAGA 1238
Qy 1353 GCTGAGGAGCGCAGCAGCTTTGGCAAGGCACACCTGAAAGCTGCGCATCTCTGGGGCAGA 1412
Db 1239 TGATCAGCAGCGCCGACAGCTGGGAAGTCCCACTTGGCAGAGCCTATGGCTCATCAG 1298
Qy 1413 GGAGGGCTGCGGCTGCAGCGCTGTGCGCGCTCTGCGGAGGTGGACCTCTCGGGC 1472
Db 1299 GGAGTCGGTGGGAGCGAGCTCTGCAAGAAAGCTGCGGAGCATTTGAGGTGCTACAGAA 1358
Qy 1473 TGTGATCTCCACAGCGCTCAGCGCTCACTGCGCAAGTATGCGGAGCTCGACCGTGGGA 1532
Db 1359 GGTGCTGTGTGCGGCACAGGAAGCTCCGGCTCAGTACGCCCGCAGCACGAGGAGGA 1418
Qy 1533 TGACTTCTGTGAGGTGCGGAGCGCCGGAATCCAGCGCTTAAGACCCACGAGAGCAGGA 1592
Db 1419 TGACTGTGAACCTGATGAGCGCCCGCAGTGTGTGTTAAATCTGAGCAGAGAGTTGA 1478
Qy 1593 GGCCAGGATGCCACGCTGTCCAGGGGAGGGGCTGACATCTTCCATTCGGCTGGGGCC 1652
Db 1479 CATTATATTGGCCCAAGTTCTCCAAGCTGACATCAAGGACTTCTTCCAGAAAGCTGGGCC 1538
Qy 1653 CCTGTCTGTGTTCTCAGCCAGAAACCGGTGGCGGTGGTGGGGCCGCTCCACCTGACCG 1712
Db 1539 CTTATCTGTGTTTTCGGCTTACAGCGGTGGAGCCTCTCGAAGCATCCGCTTCACTGC 1598
Qy 1713 AGGAGGCGCGCTTTGGCCTCACGCTTTCGGGAGACTCGCTGTCTCTCATCGCTGCGGT 1772
Db 1599 AGAAGAGGGACTTGGGGTTTCACTTCAGAGGGAACGCGCCCGCTTCAAGTTCACTTCT 1658

Qy 1773 CAUATCCAGGAGCAGCGCGCGCTGCTGAAAGGGCGGACTACATTTGTCTCAGT 1832
Db 1659 GGAATCCTTACTGCTCTGCTCGGTGGCAGGAGCCCGGAGAGATTATATTGTCTCCAT 1718
Qy 1833 GAATGGGACCCATGCAAGTGTGGAGACACGCGGAGTGTGACGAGCTGAAGGCTGC 1892
Db 1719 TCAGCTTGTGGATTGTTAAGTGGCTGACGCTGAGTGAAGTTATGAAGCTGTGAAGACTT 1778
Qy 1893 GCGAGAGCGGGCGCCAGCTGCAAGTGTGGTGTGCTGCTG 1932
Db 1779 TGGCAGGACGAGATCGAGTGAAGTCTGTGAGCCTCCTG 1818

RESULT 15
AX077672
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Yue H., Tang, Y.T., Bandman, O., Hillman, J.L., Lal, P., Au-Young, J.,
Reddy, R., Yang, J., Baughn, M.R., Lu, D.A., Azimzai, Y. and
Patterson, C.
Gtp-binding protein associated factors
Patent: WO 0105970-A 127 25-JAN-2001;
Incyte Genomics, Inc. (US)
Location/Qualifiers
1..3019
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 3357264CB1"

ORIGIN

Query Match 22.5%; Score 487.8; DB 6; Length 3019;
Best Local Similarity 56.9%; Pred. No. 1.3e-66;
Matches 957; Conservative 0; Mismatches 713; Indels 12; Gaps 3;
Qy 257 TGCAGAGCGCAGGCGCCAGATTCCACGACGATTCACAGGAGCTGCAGATGCGGACCG 316
Db 190 TGCAGATCAAGAGCTGCTTTGAATCAGCATCTCTGAAAGCCGTGCGGATGAGGACCG 249
Qy 317 GCGCTGAGAACCTCTACAGAGCCACGACCAACCCGGTGAGAGACGCTGCGCCTGG 376
Db 250 GAGCGAAACCTTCTGAAAGTGGCCACAACTCAAGGTGCGGAGCAAGTGGCGCTGG 309
Qy 377 AGCTGAGCTAGCTCAACTCCAACTGAGCTGCTGCAAGGAGCTGAGGAGCTCAAGC 436
Db 310 AGCTGAGCTTGTCAACTCAGACCTGCGATGCTCAAGGAGAGCTGAGGAGCTGAACA 369
Qy 437 GTGGCTGAGCCCTCGGCGCATGGAGAGCTGCTCACTGCTCCCATGATCCCGCTGG 496
Db 370 TCTCGTGGGCTGTATCAGAAACAGAGGAGGAGTTTACGATTCCTCGATTCCTCTTG 429
Qy 497 GCTTGAAGAGACCAAGAGCTGGACTGTCTACACCCCTGAAGGAGCTGATCTCAGTGC 556
Db 430 GCTTGAAGAAACGAAAGACGCTGACATTGTCAGTCTCTCAAGGATTTTATCTGGAAAC 489
Qy 557 ACTTGGAGAGAGCGGCCCTCTACGAGGAGAAATCAGGAGCTGGAGCCCTTGGCGC 616
Db 490 ATTAGTGAAGATGGCTATTATATGAAGTGAATTCAGATCTTATGATGATCTGAGAC 549
Qy 617 AGGCATGCGGACCCCGAGCGGATGAGTGGGCGCTGGAGCTGTACAGCCTATTACA 676
Db 550 AAGCTTGTGGAGCCCTAGCCGGATGAGCGCGGGTGGAACTGTGATGACATCTTCA 609
Qy 677 ACCAGCTGTGCTTCTCGATGCGGCTTCTCTCACCCCTGCCAGGAGCCTCGGGCTCTTCT 736

Db	610	TCACGCTGGGCTTTGTCGAGAGTCGATTTCTCCGGCCACACGCGCAGATGGGACTCTCTGT	669
Qy	737	TCCACTGGTACGACTCGCTTACTGGGTCCCGGCCCCAGCAGCGCTGCCCCCTGCCCCCTCGAGA	796
Db	670	TCACCTGGTATGACTCTCTCACCGGGTTCGGGTACGCCAGCAGAACTCTGCTGAGAGA	729
Qy	797	AGGGCAGCGTCTCTTCAACATCGGTGCCCTCAACAGCAGATTTGGGGCGCGCCAGGACC	856
Db	730	AGGCCAGTGTCTGTTCAACACTGGGGCCCTCTACACCCAGATTTGGGACCCGGTGTGATC	789
Qy	857	GCTCCTGCACCGAGGGTGCCCGCGCTATGAGGCTTCCAGAGGGCGCTGGGGCCT	916
Db	790	GGCAGACGCGAGGCTGGGCTGGAGAGTGCCATAGATGCTTTTCAGAGCGCGCAGGGTTT	849
Qy	917	TCAGCCTTCCTGAGGGAGAACTTCTCCCATGCGCGAGCCCGCAGACATAGCGCTGCGTCCC	976
Db	850	TAAATTACCTGAAAGACACATTTACCCATCTCCAAAGTTACGACATGAGCCCTGCCATGC	909
Qy	977	TCTGCGCAGCTGGAGCAGCTCATGATGGCCCGAGGCCCAGGAATGTGTGTTTGGGGCCTCT	1036
Db	910	TCAGCGTGTCTCGTCAAAATGATGCTTGCAAGCCCAAGAACGCGTGTTTGAGAAATCA	969
Qy	1037	CACCACTGCTCCATGCTGCCCCCAAGACTGCTGCGCCAGCTGCGCTGGCGCAGGAGG	1096
Db	970	GCCCTTCCTG-----GGATCCNGAATGAATCTTCATGCTGGTGAAGGTGGCTCAGGAGG	1023
Qy	1097	CCGCCAGGTGGCAGCGAGTACAGGCTAGTGCACCGGACCATGGCCGAGCACCCCGTCC	1156
Db	1024	CTGCTAAGGTGGAGAGTCTTACCAACAGCTACACGAGCCATGAGCCAGGCGCGGTGA	1083
Qy	1157	ACGACTAGTGCCTGTCTCTGGACTGCCCTGGTGCATGTCAAGGCCGAGTACTTCCGCT	1216
Db	1084	AAGAGAACATCCCTTACTCTCTGGGCCAGCTTAGCGTGGTGAAGGCCCAACACTACGGGG	1143
Qy	1217	CCCTGGCCCACTACACGTAGCCATGGCCCTCTGCAAGCGCTCCCCAGCGACCCGAGGGAG	1276
Db	1144	CCCTGGCCCACTACTTCTACTGCCATCTTCTCTCATGACCACAGGTGAAGCCAGGCACGG	1203
Qy	1277	AGCTCCCCCAGCAGCAGAGTCTTCTGTCAGC-----CCCCCACTCTCTTAAGCCCCGA	1332
Db	1204	ATCTGGACCACAGGAGAGTGCCCTGTCCAGCTCTAGAGCCACATGCCAGAGGGGTGA	1263
Qy	1333	GGCCCTGTGC--TGCCGAGAGCTGGAGAGCGCAGCAGCTTGGCAAGGCACACCTGA	1390
Db	1264	CACCCCTGGCCACACTGAAGAATGATCAGCAGCGCCGACAGCTGGGGAGTCCCACTTGC	1323
Qy	1391	AGCGTGCCATCTGGGGCAGGAGGCGCTGCGGCTGCACGCCCTGTGCGCGCTCTTC	1450
Db	1324	GCAGAGCCATGGCTCATCAGAGGAGTCGGTGGGGGCGAGGCTCTGCAAGAAGCTGC	1383
Qy	1451	GGAGGTGGACCTGTTGGGGCTGTGATCTCCAGACGCTGAGCGCTCACTGGGCCAAGT	1510
Db	1384	GGACGATTTGAGGTGCTACAGAAGTGTCTGTGCGCCACAGGAACGCTCCCGGCTCACGT	1443
Qy	1511	ATGCGGAGCTGCACGCTGAGGATGACTTCTGTGAGGCTGCGAGGCCCGGACATCCAGC	1570
Db	1444	AGCCCGACGACAGGAGGATGACCTGCTGAACCTGATCAGCGCCCCCAGTGTGTG	1503
Qy	1571	CTAAGACCCACACAGAAGCCAGAGGCCAGGATGCCACGCTGTCCAGGGGAAAGGGCCCTG	1630
Db	1504	CTAAGACTGAGCAAGAGTTGACATTAATTTGCCCCAGTTCTCCAAAGCTGACAGTCA	1563
Qy	1631	ACATCTTCATCGGCTGGGGCCCTGTCTGTGTTCTCAGCCAAAGAACCGGTGGCGGCTGG	1690
Db	1564	ACTTCTTCAGAAAGCTGGGCCCTTATCTGTGTTTTCGGCTAACCAAGCGGTGGACGCTTC	1623
Qy	1691	TGGGGCCCTGCACCTGACCGAGGAGAGGGCGGCTTTGGCTCTCAGCTTCGGGGAGACT	1750
Db	1624	CTCGAAGCATCCGCTTCACTGCAGAAGAAGGGGACTTTGGGTTTCACTTGAAGGGAA	1683
Qy	1751	CGCCTGTCTCATCGTCTCGTCAATTCAGGAGGCGCGCGCGCTGGCCTGAAGG	1810

Search completed: July 13, 2004, 06:10:05
Job time : 5576.17 secs